

STIC-Biotech/ChemLib

145922

me

From: Swope, Sheridan
Sent: Tuesday, February 22, 2005 4:59 PM
To: STIC-Biotech/ChemLib
Subject: 10/018,964

For 10/018,964, pls search and interference search:

SID 3, residues 1-15, against the NT and AA data bases.

Please separate out the interference searches.

Sheridan Swope, Ph.D.
Patent Examiner, AU 1652
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E02C70 Remsen Bld (Mailbox)

RECEIVED
FEB 23 2005
STIC

Barb O'Brien

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # _____
AA Sequence: # _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2005, 13:50:51 ; Search time 170 Seconds
(without alignments)
45.183 Million cell updates/sec

Title: US-10-018-964-3_COPY_1_15
Perfect score: 87
Sequence: 1 YSGPPSGARRNCYE 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	100.0	319	1 MYOD_HUMAN	P15172 homo sapien
2	87	100.0	320	2 O75321	O75321 homo sapien
3	84	96.6	150	2 O77799	O77799 bos taurus
4	84	96.6	318	2 O7YS82	O7YS82 bos taurus
5	84	96.6	319	1 MYOD_PIG	P49811 sus scrofa
6	84	96.6	319	1 MYOD_SHEEP	P29331 ovis aries
7	63	72.4	318	1 MYOD_MOUSE	P10085 mus musculus
8	63	72.4	318	1 MYOD_RAT	Q02346 rattus norv
9	63	72.4	318	2 O8C6B1	O8C6B1 mus musculus
10	60	69.0	297	1 MYOD_CONJA	P21572 coturnix co
11	60	69.0	297	2 Q6DV59	Q6DV59 meleagris g
12	60	69.0	298	1 MYOD_CHICK	P16075 gallus gall
13	57	65.5	289	1 MYOD_XENLA	P13904 xenopus lae
14	57	65.5	289	2 O8AVZ0	O8AVZ0 xenopus lae
15	50	57.5	172	2 O7SV41	O7SV41 physcomitre
16	50	57.5	417	2 O7SV42	O7SV42 physcomitre
17	49.5	56.9	288	2 Q7T109	Q7T109 xenopus tro
18	48	55.2	652	2 Q8WSN9	Q8WSN9 dugesia tig
19	48	55.2	652	2 Q8WSX0	Q8WSX0 dugesia tig
20	47	54.0	271	1 FRAL_HUMAN	P148755 mus sapien
21	47	54.0	273	1 FRAL_MOUSE	P48755 mus musculus
22	47	54.0	273	2 Q7TMU4	Q7TMU4 mus musculus
23	47	54.0	275	1 FRAL_RAT	P10158 rattus norv
24	47	54.0	652	2 Q8WSX1	Q8WSX1 dugesia tig
25	47	54.0	2340	2 Q7RRD0	Q7RRD0 plasmodium
26	46	52.9	932	2 Q9U966	Q9U966 eimeria ten
27	45.5	52.3	313	2 Q7XE93	Q7XE93 oryza sativ
28	45	51.7	120	2 Q9ZWY5	Q9ZWY5 bacterioph
29	45	51.7	139	2 Q6ZAR3	Q6ZAR3 oryza sativ
30	45	51.7	1184	1 DP3A_MYCBO	P63978 mycobacteri
31	45	51.7	1184	1 DP3A_MYCTU	P63977 mycobacteri

Q05400 streptomyce
O7q307 anopheles g
O7z9g9 brachydanio
P16076 xenopus lae
Q8gn48 xenopus lae
Q84w52 arabidopsis
Q91fb9 arabidopsis
Q9a9j5 caulobacter
Q8pay6 xanthomonas
Q8lhg2 photobacter
Q8mm82 caenorhabdi
Q8gj48 chlorobium
Q8gj49 chlorobium
Q8gj35 chlorobium

ALIGNMENTS

RESULT 1

MYOD_HUMAN ID MYOD_HUMAN STANDARD; PRT; 319 AA.
AC P15172;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Myoblast determination protein 1 (Myogenic factor 3) (Myf-3).
GN Name=MYOD1; Synonyms=MYF3, MYOD;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91212196; PubMed=1850513;
RA Pearson-White S.H.;
RT "Human MyoD: cDNA and deduced amino acid sequence."
RL Nucleic Acids Res. 19:1148-1148(1991).
RN [2]
RP SEQUENCE OF 53-319 FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=90059960; PubMed=2583111;
RA Braun T., Bober E., Buschhausen-Denker G., Kotz S., Grzeschik K.-H.,
RA Arnold H.H.;
RT "Differential expression of myogenic determination genes in muscle
cells: possible autoactivation by the Myf gene products."
RL EMBO J. 8:3617-3625(1989).
RN [3]
RP REVIEW ABOUT ACETYLATION AND DEACETYLATION.
RX MEDLINE=21423399; PubMed=11532390; DOI=10.1016/S0959-437X(00)00224-0;
RA McKinsey T.A., Zhang C.L., Olson E.N.;
RT "Control of muscle development by dueling HATs and HDACs."
RL Curr. Opin. Genet. Dev. 11:497-504(2001).
CC -I- FUNCTION: Involved in muscle differentiation (myogenic factor).
CC Induces fibroblasts to differentiate into myoblasts. Activates
CC muscle-specific promoters. Interacts with and is inhibited by the
CC twist protein. This interaction probably involves the basic
CC domains of both proteins (By similarity).
CC -I- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein. Seems to form active heterodimers with Irf-2.
CC -I- SUBCELLULAR LOCATION: Nuclear.
CC -I- PTM: Acetylated by a complex containing EP300 and PCAF. The
CC acetylation is essential to activate target genes. Conversely, its
CC deacetylation by SIRT1 inhibits its function (By similarity).
CC -I- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL; X56677; CAA40000.1; -.
DR EMBL; X17650; CAA35640.1; -.
DR FIR; Z26827; Z26827.
DR HSSP; P10085; 1MDY.
DR TRANSFAC; T00519; -.
DR TRANSFAC; T00525; -.
DR Genew; HGNC:7611; MYOD1.
DR MIM; 159970; -.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0003705; F:RNA polymerase II transcription factor acti. .; TAS.
DR GO; GO:0003713; F:transcription coactivator activity; TAS.
DR GO; GO:0007519; P:myogenesis; TAS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR GO; GO:0006357; P:regulation of transcription from Pol II pro. .; TAS.
DR InterPro; IPR002546; Basic.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF01586; Basic; 1.
DR Pfam; PF00010; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
KW Acetylation; Developmental protein; Differentiation; DNA-binding;
KW Myogenesis; Nuclear protein; Transcription regulation.
FT DNA BIND 109 121
FT DOMAIN 122 161 Helix-loop-helix motif.
FT CONFLICT 124 124 K -> E (in Ref. 2).
SQ SEQUENCE 319 AA; 34490 MW; AAE935C154318770 CRC64;

Query Match 100.0%; Score 87; DB 1; Length 319;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSGPPSGARRRNCYE 15
DB 212 YSGPPSGARRRNCYE 226
|||||

RESULT 2
ID 075321 PRELIMINARY; PRT; 320 AA.
AC 075321;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Myogenic determining factor 3 (Myogenic factor 3).
GN Name=MYOD1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9820644; PubMed=9546368;
RA Chen B., Dias P., Jenkins J.J. III, Savell V.H., Parham D.M.;
RT "Methylation alterations of the MyoD1 upstream region are predictive
RL of subclassification of human rhabdomyosarcomas.";
RL Am. J. Pathol. 152:1071-1079(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA Phelan M., Farmer A.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.R.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Scheetz C.,

```

```

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC -!- BHLH protein (by similarity).
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
DR EMBL; AF027148; AAC29001.1; -.
DR EMBL; BT007157; AAP35821.1; -.
DR EMBL; BC064493; AAG64493.1; -.
DR HSSP; P10085; 1MDY.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0007517; P:muscle development; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR002546; Basic.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF01586; Basic; 1.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00520; BASIC; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 320 AA; 34501 MW; 75E624D2ED5B0833 CRC64;

Query Match 100.0%; Score 87; DB 2; Length 320;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSGPPSGARRRNCYE 15
DB 213 YSGPPSGARRRNCYE 227
|||||

RESULT 3
ID 077799 PRELIMINARY; PRT; 150 AA.
AC 077799;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Myogenic regulatory factor (Fragment).
GN Name=MyoD;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21193248; PubMed=11294773;
RA Oldham J.M., Martyn J.A., Sharma M., Jeanplong F., Kanbadur R.,
RA Bass J.J.;
RT "Molecular expression of myostatin and MyoD is greater in double-
RT muscled than normal-muscled cattle fetuses";
RL Am. J. Physiol. Regul. Integr. Comp. Physiol. 280:R1488-R1493(2001).
DR EMBL; AF093675; AAC63373.1; -.
FT NON_TER 1
FT NON_TER 150 150

```



```

SQ SEQUENCE 150 AA; 15226 MW; 0C62897AC3D6A477 CRC64;
Query Match 96.6%; Score 84; DB 2; Length 150;
Best Local Similarity 93.3%; Pred. No. 2.2e-05;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YSGPPSGARRNCYE 15
Db 45 YSGPPSGARRNCYD 59

RESULT 4
Q7YS82 PRELIMINARY; PRT; 318 AA.
AC Q7YS82
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE MYOD.
GN Bos taurus (Bovine).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RA Muroya S., Nakajima I., Chikuni K.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein (By similarity).
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
DR EMBL: AB110599; BAC76802.1; -.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0007517; P:muscle development; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR002546; Basic.
DR InterPro: IPR001092; HLH_basic.
DR Pfam: PF01586; Basic; 1.
DR SMART: SM00520; BASIC; 1.
DR SMART: SM00353; HLH; 1.
DR PROSITE: PS00888; HLH; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 318 AA; 34207 MW; BD467B506D0A5447 CRC64;

Query Match 96.6%; Score 84; DB 2; Length 318;
Best Local Similarity 93.3%; Pred. No. 4.7e-05;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YSGPPSGARRNCYE 15
Db 213 YSGPPSGARRNCYD 227

RESULT 5
MYOD_PIG
ID MYOD_PIG STANDARD; PRT; 319 AA.
AC P49811;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Myoblast determination protein 1.
GN Name=MYOD1; Synonyms=MYOD;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Large white X Landrace;

RX MEDLINE=96024096; PubMed=7559997;
RA Chang K.C., Fernandes K., Chantler P.D.;
RT "Cloning and in vivo expression of the pig MyoD gene.";
RL J. Muscle Res. Cell Motil. 16:243-247(1995).
CC -1- FUNCTION: Involved in muscle differentiation (myogenic factor).
CC Induces fibroblasts to differentiate into myoblasts. Activates
CC muscle-specific promoters. Interacts with and is inhibited by the
CC twist protein. This interaction probably involves the basic
CC domains of both proteins (By similarity).
CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein. Seems to form active heterodimers with ITF-2 (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: Acetylated by a complex containing EP300 and PCAF. The
CC acetylation is essential to activate target genes. Conversely, its
CC deacetylation by SIR1 inhibits its function (By similarity).
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: U12574; AAA87061.1; -.
CC HSSP: P10085; LMDY.
CC InterPro: IPR002546; Basic.
CC InterPro: IPR001092; HLH_basic.
CC Pfam: PF01586; Basic; 1.
CC Pfam: PF00010; HLH; 1.
CC PROSITE: PS00888; HLH; 1.
CC ACetylation; Developmental protein; Differentiation; DNA-binding;
CC Myogenesis; Nuclear protein; Transcription regulation.
CC DNA BIND 109 121 Basic motif.
FT DOMAIN 122 161 Helix-loop-helix motif.
SQ SEQUENCE 319 AA; 34243 MW; B84E5A5C8CF9B288 CRC64;

Query Match 96.6%; Score 84; DB 1; Length 319;
Best Local Similarity 93.3%; Pred. No. 4.7e-05;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YSGPPSGARRNCYE 15
Db 213 YSGPPSGARRNCYD 227

RESULT 6
MYOD_SHEEP
ID MYOD_SHEEP STANDARD; PRT; 319 AA.
AC P29331;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Myoblast determination protein 1.
GN Name=MYOD1; Synonyms=MYOD;
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Coopworth; TISSUE=Thigh muscle;
RX MEDLINE=92158635; PubMed=1741269;
RA Huynen L., Bass J., Gardner R.C., Bellamy A.R.;
RT "Nucleotide sequence of the sheep MyoD1 gene.";
RL Nucleic Acids Res. 20:374-374(1992).
CC -1- FUNCTION: Involved in muscle differentiation (myogenic factor).
CC Induces fibroblasts to differentiate into myoblasts. Activates
CC muscle-specific promoters. Interacts with and is inhibited by the
CC twist protein. This interaction probably involves the basic

```

```

CC domains of both proteins (By similarity).
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC BHLH protein. Seems to form active heterodimers with ITP-2 (By
CC similarity).
CC
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC
CC -!- PTM: Acetylated by a complex containing EP300 and PCAF. The
CC acetylation is essential to activate target genes. Conversely, its
CC deacetylation by SIRT1 inhibits its function (By similarity).
CC
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X62102; CAA44012.1; -
CC PIR; S20086; S20086.
CC HSP; P10085; IMDY.
CC InterPro; IPR002546; Basic.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF01586; Basic; 1.
CC Pfam; PF00010; HLH; 1.
CC PROSITE; PS50888; HLH; 1.
CC Acetylation; Developmental protein; Differentiation; DNA-binding;
KW Myogenesis; Nuclear protein; Transcription regulation.
FT DNA_BIND 109 121 Basic motif.
FT DOMAIN 122 161 Helix-loop-helix motif.
SQ SEQUENCE 319 AA; 34209 MW; 92ECB2BAD568E0D CRC64;

Query Match 96.6%; Score 84; DB 1; Length 319;
Best Local Similarity 93.3%; Pred. No. 4.7e-05;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSGPPSGARRRNCYE 15
Db 213 YSGPPSGARRRNCYD 227
|||||
|||||

RESULT 7
MYOD MOUSE STANDARD; PRT; 318 AA.
AC P10085;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Myoblast determination protein 1.
GN Name=Myod1; Synonyms=Myod;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88080485; PubMed=3690668; DOI=10.1016/0092-8674(87)90585-X;
RA Davis R.L., Weintraub H., Lassar A.B.;
RT "Expression of a single transfected cDNA converts fibroblasts to
RT myoblasts."
RL Cell 51:987-1000(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92093599; PubMed=1754380;
RA Zingg J.M., Alva G.P., Joet J.P.;
RT "Characterisation of a genomic clone covering the structural mouse
RT MyoD1 gene and its promoter region."
RL Nucleic Acids Res. 19:6433-6439(1991).
RN [3]
RP FUNCTION.
RX MEDLINE=89019371; PubMed=3175662;
RA Tapscott S.J., Davis R.L., Thayer M.J., Cheng P.-F., Weintraub H.,
RA Lassar A.B.;

```

```

RT "MyoD1: a nuclear phosphoprotein requiring a Myc homology region to
RT convert fibroblasts to myoblasts."
RL Science 242:405-411(1988).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=88223371; PubMed=3286015; DOI=10.1016/0092-8674(88)90095-5;
RX Pinney D.F., Pearson-White S.H., Konieczny S.F., Latham K.E.,
RA Emerson C.P. Jr.;
RT "Myogenic lineage determination and differentiation: evidence for a
RT regulatory gene pathway."
RL Cell 53:781-793(1988).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 102-166.
RX MEDLINE=94236689; PubMed=8181063; DOI=10.1016/0092-8674(94)90159-7;
RX Ma P.C.M., Rould M.A., Weintraub H., Pabo C.O.;
RA "Crystal structure of MyoD bHLH domain-DNA complex: perspectives on
RA DNA recognition and implications for transcriptional activation."
RL Cell 77:451-459(1994).
RN [6]
RP INHIBITION BY TWIST.
RX MEDLINE=98001585; PubMed=9343420;
RX Hamamori Y., Wu H.Y., Sartorelli V., Keddes L.;
RA "The basic domain of myogenic basic helix-loop-helix (bHLH) proteins
RA is the novel target for direct inhibition by another bHLH protein,
RA Twist."
RL Mol. Cell. Biol. 17:6563-6573(1997).
RN [7]
RP ACETYLATION.
RX MEDLINE=97180938; PubMed=9029156; DOI=10.1093/emboj/16.2.369;
RX Puri P.L., Avantaggiati M.L., Balsano C., Sang N., Graessmann A.,
RA Giordano A., Leviero M.;
RA "p300 is required for MyoD-dependent cell cycle arrest and muscle-
RA specific gene transcription."
RL EMBO J. 16:369-383(1997).
RN [8]
RP DEACETYLATION BY SIRT1.
RX MEDLINE=22770674; PubMed=12887892; DOI=10.1016/S1097-2765(03)00226-0;
RX Fulco M., Schiltz R.L., Iezzi S., King M.T., Zhao P., Kashiwaya Y.,
RA Hoffman E., Veech R.L., Sartorelli V.;
RT "Sir2 regulates skeletal muscle differentiation as a potential sensor
RT of the redox state."
RL Mol. Cell 12:51-62(2003).
RN [9]
RP REVIEW ABOUT ACETYLATION AND DEACETYLATION.
RX MEDLINE=21423399; PubMed=11532390; DOI=10.1016/S0959-437X(00)00224-0;
RX McKinsey T.A., Zhang C.L., Olson E.N.;
RT "Control of muscle development by dueling HATs and HDACs."
RL Curr. Opin. Genet. Dev. 11:497-504(2001).
CC -!- FUNCTION: Involved in muscle differentiation (myogenic factor).
CC Induces fibroblasts to differentiate into myoblasts. Activates
CC muscle-specific promoters. Interacts with and is inhibited by the
CC twist protein. This interaction probably involves the basic
CC domains of both proteins.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein. Seems to form active heterodimers with ITP-2.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- PTM: Acetylated by a complex containing EP300 and PCAF. The
CC acetylation is essential to activate target genes. Conversely, its
CC deacetylation by SIRT1 inhibits its function.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M18779; AAA39799.1; -
CC EMBL; X61655; CAA43836.1; -
CC EMBL; M84918; AAA39798.1; -
CC PIR; A29636; A29636.

```

```
DR PDB, INDY; X-ray; A/B/C/D=-.
DR TRANSFAC; T00526; -.
DR MGI; MGI:97275; Myod1.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR002546; Basic.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF01586; Basic; 1.
DR PROSITE; PS0888; HLH; 1.
DR PROSITE; PS0888; HLH; 1.
DR 3D-structure; Acetylation; Developmental protein; Differentiation;
KW DNA-binding; Myogenesis; Nuclear protein; Transcription regulation.
FT DNA_BIND 109 121
FT DOMAIN 122 161
FT HELIX 122 161
FT TURN 134 135
FT TURN 140 141
FT TURN 146 162
FT TURN 163 164
SQ SEQUENCE 318 AA; 34218 MW; 48C836896FCBDC0C CRC64;

Query Match 72.4%; Score 63; DB 1; Length 318;
Best Local Similarity 73.3%; Pred. No. 0.096;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YSGPPSGARRRNCYE 15
Db 212 YSGPPSGRRQRNGYD 226

RESULT 8
MYOD_RAT
ID MYOD_RAT STANDARD; PRT; 318 AA.
AC Q02346;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Myoblast determination protein 1.
GN Name=Myod1; Synonyms=Myod;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92339895; PubMed=1321778; DOI=10.1016/0378-1119(92)90519-U;
RA Vaidya T.B., Rhodes S.J., Moore J.L., Sherman D.A., Konieczny S.F.,
RA Taparowsky E.J.;
RT "Isolation and structural analysis of the rat Myod gene.";
RL Gene 116:223-230(1992).
CC -1- FUNCTION: Involved in muscle differentiation (myogenic factor).
CC Induces fibroblasts to differentiate into myoblasts. Activates
CC muscle-specific promoters. Interacts with and is inhibited by the
CC twist protein. This interaction probably involves the basic
CC domains of both proteins (By similarity).
CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein. Seems to form active heterodimers with IRF-2 (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: Acetylated by a complex containing EP300 and PCAF. The
CC acetylation is essential to activate target genes. Conversely, its
CC deacetylation by SIR1 inhibits its function (By similarity).
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M84176; AAA41661.1; -.
```

```
DR PIR; JCI171; JCI171.
DR HSSP; P10085; IMDY.
DR InterPro; IPR002546; Basic.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF01586; Basic; 1.
DR Pfam; PF00010; HLH; 1.
DR PROSITE; PS0888; HLH; 1.
DR PROSITE; PS0888; HLH; 1.
KW Acetylation; Developmental protein; Differentiation; DNA-binding;
KW Myogenesis; Nuclear protein; Transcription regulation.
FT DNA_BIND 109 121
FT DOMAIN 122 161
FT SEQUENCE 318 AA; 34359 MW; FC43E468B1287F5F CRC64;

Query Match 72.4%; Score 63; DB 1; Length 318;
Best Local Similarity 73.3%; Pred. No. 0.096;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YSGPPSGARRRNCYE 15
Db 212 YSGPPSGRRQRNGYD 226

RESULT 9
Q8C6B1
ID Q8C6B1 PRELIMINARY; PRT; 318 AA.
AC Q8C6B1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length enriched
DE library, clone:320002D16 product:myogenic differentiation 1, full
DE insert sequence.
GN Name=Myod1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Head;
RA The RIKEN Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
```


DR SMART; SM00520; BASIC; 1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS0888; HLH; 1.
 KW DNA-binding; Nuclear protein.
 SQ SEQUENCE 297 AA; 32906 MW; CD7D16B29EFD5243 CRC64;
 Query Match 69.0%; Score 60; DB 2; Length 297;
 Best Local Similarity 66.7%; Pred. NO. 0.27;
 Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 YSGPPSGARRRNCYE 15
 Db 190 YSGPPCSRRRNSYD 204
 RESULT 12
 MYOD CHICK
 ID MYOD CHICK STANDARD; PRT; 298 AA.
 AC P16075; Q90916;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Myoblast determination protein 1 homolog (MYOD1 homolog).
 GN Name=MYOD1; Synonyms=CMD1, MYOD;
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89378745; PubMed=2777078;
 RA Lin Z.Y., Dechesne C.A., Eldridge J., Patterson B.M.;
 RT "An avian muscle factor related to MyoD1 activates muscle-specific
 RT promoters in nonmuscle cells of different germ-layer origin and in
 RT BrdU-treated myoblasts.";
 RL Genes Dev. 3:986-996(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPAFAS; TISSUE=Skeletal muscle;
 RX MEDLINE=94309668; PubMed=8035824;
 RA Dechesne C.A., Wei Q., Eldridge J., Gannoun-Zaki L., Millasseau P.,
 RA Bougueleret L., Caterina D., Patterson B.M.;
 RT "B-box- and MEF-2-independent muscle-specific expression, positive
 RT autoregulation, and cross-activation of the chicken MyoD (CMD1)
 RT promoter reveal an indirect regulatory pathway.";
 RL Mol. Cell. Biol. 14:5474-5486(1994).
 CC -1- FUNCTION: Involved in muscle differentiation (myogenic factor).
 CC Activates muscle-specific promoters. Interacts with and is
 CC inhibited by the twist protein. This interaction probably involves
 CC the basic domains of both proteins (By similarity).
 CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein. Seems to form active heterodimers with ITF-2.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X16189; CAA34315.1; -;
 CC EMBL; L34006; AAA74374.1; -;
 CC PIR; A32872; A32872.
 CC HGSP; P10085; 1MDY.
 CC TRANSFAC; T01128; -;
 CC InterPro; IPR002546; Basic.
 CC InterPro; IPR001092; HLH_Basic.
 CC Pfam; PF01586; Basic; 1.
 CC Pfam; PF00010; HLH; 1.

DR SMART; SM00520; BASIC; 1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS0888; HLH; 1.
 KW Developmental protein; Differentiation; DNA-binding; Myogenesis;
 KW Nuclear protein; Transcription regulation.
 FT DNA BIND 100 112 Basic motif.
 FT DOMAIN 113 152 Helix-loop-helix motif.
 FT CONFLICT 53 53 P -> A (in Ref. 1).
 SQ SEQUENCE 298 AA; 32990 MW; 1C001869B4657B7A CRC64;
 Query Match 69.0%; Score 60; DB 1; Length 298;
 Best Local Similarity 66.7%; Pred. NO. 0.27;
 Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 YSGPPSGARRRNCYE 15
 Db 189 YSGPPCSRRRNSYD 203
 RESULT 13
 MYOD XENLA
 ID MYOD XENLA STANDARD; PRT; 289 AA.
 AC P13904;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Myoblast determination protein 1 homolog (Myogenic factor 1).
 GN Name=MYOD; Synonyms=MFI, clawed frog).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90059936; PubMed=2555164;
 RA Hopwood N.D., Pluck A., Gurdon J.B.;
 RT "MyoD expression in the forming somites is an early response to
 RT mesoderm induction in Xenopus embryos.";
 RL EMBO J. 8:3409-3417(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90205830; PubMed=1690844;
 RA Scales J.B., Olson E.N., Perry M.;
 RT "Two distinct Xenopus genes with homology to MyoD1 are expressed
 RT before somite formation in early embryogenesis.";
 RL Mol. Cell. Biol. 10:1516-1524(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90353186; PubMed=2167198;
 RA Harvey R.P.;
 RT "The Xenopus MyoD gene: an unlocalised maternal mRNA predates lineage-
 RT restricted expression in the early embryo.";
 RL Development 108:669-680(1990).
 CC -1- FUNCTION: Myogenesis.
 CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- DEVELOPMENTAL STAGE: Expression is specific to the mesoderm of the
 CC gastrula.
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X16106; CAA34232.1; -;
 CC EMBL; M31116; AAA49900.1; -;
 CC PIR; A34783; A34783.

```
DR HSSP; P10085; IMDY.
DR TRANSFAC; T00524; -.
DR InterPro; IPR002546; Basic.
DR Pfam; PF01586; Basic; 1.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00520; BASIC; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
KW Developmental protein; Differentiation; DNA-binding; Myogenesis;
KW Nuclear protein; Transcription regulation.
FT DNA_BIND 95 107
FT DOMAIN 108 147
FT CONFLICT 255 256
SQ SEQUENCE 289 AA; 32306 MW; D0A985DC66136284 CRC64;

Query Match 65.5%; Score 57; DB 1; Length 289;
Best Local Similarity 60.0%; Pred. No. 0.77;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 YSGPPSGARRNCYE 15
Db 184 YNSPPCGSRRRNSYD 198

RESULT 14
Q8AVZ0 PRELIMINARY; PRT; 289 AA.
AC Q8AVZ0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE MGCS2596 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RP SEQUENCE FROM N.A.
```

```
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein (by similarity).
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
DR EMBL; BC041190; AAH41190.1; -.
DR HSSP; P10085; IMDY.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0007517; P:muscle development; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR002546; Basic.
DR InterPro; IPR001092; HLH_Basic.
DR Pfam; PF01586; Basic; 1.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00520; BASIC; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 289 AA; 32322 MW; 714985C267FBB972 CRC64;

Query Match 65.5%; Score 57; DB 2; Length 289;
Best Local Similarity 60.0%; Pred. No. 0.77;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 YSGPPSGARRNCYE 15
Db 184 YNSPPCGSRRRNSYD 198

RESULT 15
Q75V41 PRELIMINARY; PRT; 172 AA.
AC Q75V41;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Plastid division site determinant Mind (Fragment).
GN Name=PpMind2;
OS Physcomitrella patens (Moss).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
OX NCBI_TaxID=3218;
RN [1];
RP SEQUENCE FROM N.A.
RA Itoh R., Machida M., Takechi K., Takio S., Fujita T., Hasebe M.,
RA Takano H.;
RT "Physcomitrella patens gene for mind2 (PpMind2).";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB125339; BAD18010.1; -.
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 18049 MW; 1B68C6F15D4EB5B9 CRC64;

Query Match 57.5%; Score 50; DB 2; Length 172;
Best Local Similarity 61.5%; Pred. No. 5.9;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SGPPSGARRNCY 14
Db 5 TGPPSGALQNGCY 17

Search completed: March 3, 2005, 14:00:07
Job time : 171 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2005, 13:42:35 ; Search time 39 Seconds
(without alignments)
37,006 Million cell updates/sec

Title: US-10-018-964-3_COPY_1_15

Perfect score: 87

Sequence: 1 YSGPPSGARRNCYE 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	100.0	319	2	S26827
2	84	96.6	319	2	S20086
3	63	72.4	318	2	A29636
4	63	72.4	318	2	JC1171
5	60	69.0	297	2	A35874
6	60	69.0	298	2	A32872
7	57	65.5	289	2	A34783
8	47	54.0	271	2	S15750
9	47	54.0	275	1	VRTRFR
10	45	51.7	1184	2	H70761
11	44	50.6	127	2	S11911
12	44	50.6	287	2	C34783
13	44	50.6	568	2	T45957
14	44	50.6	949	2	E87371
15	43	49.4	184	2	T31322
16	43	49.4	346	2	S36665
17	43	49.4	376	2	T48714
18	43	49.4	633	2	H84854
19	43	49.4	809	2	T20430
20	43	49.4	1122	2	T47424
21	41	47.1	184	2	B87708
22	41	47.1	184	2	D87376
23	41	47.1	233	2	I65197
24	41	47.1	266	2	B72532
25	41	47.1	372	2	T28027
26	41	47.1	380	2	T28251
27	40.5	46.6	321	2	E97314
28	40.5	46.6	328	2	S35336
29	40.5	46.6	391	2	B87449

30	40.5	46.6	597	2	T47673
31	40	46.0	122	2	A87295
32	40	46.0	148	1	QOLJX2
33	40	46.0	275	2	S47325
34	40	46.0	376	2	A72389
35	40	46.0	471	2	E86322
36	40	46.0	506	2	G90000
37	40	46.0	686	2	S60749
38	40	46.0	807	2	S51460
39	40	46.0	1310	1	W2BE62
40	40	46.0	4135	2	T42629
41	39	44.8	28	2	A41476
42	39	44.8	144	2	A85440
43	39	44.8	198	2	T26708
44	39	44.8	219	1	A39574
45	39	44.8	251	2	C97437

ALIGNMENTS

RESULT 1

S26827

myogenic factor 3 - human

N:Alternate names: myoD protein; myogenic determination factor Myf-3

C:Species: Homo sapiens (man)

C:Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 09-Jul-2004

C:Accession: S26827; S06947

R:Pearson-White, S.H.

Nucleic Acids Res. 19, 1148, 1991

A:Title: Human MyoD: cDNA and deduced amino acid sequence.

A:Reference number: S26827; MUID:91212198; PMID:1850513

A:Accession: S26827

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-319 <PEA>

A:Cross-references: UNIPROT:P15172; EMBL:X56677; NID:G34861; PIDN:CRA40000.1; PID:G34862

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1990

R:Braun, T.; Bober, E.; Buschhausen-Denker, G.; Kotz, S.; Grzeschik, K.H.; Arnold, H.H.

EMBO J. 8, 3617-3625, 1989

A:Title: Differential expression of myogenic determination genes in muscle cells: possib

A:Reference number: S06947; MUID:90059960; PMID:2583111

A:Accession: S06947

A:Molecule type: mRNA

A:Residues: 53-123, 'E', 125-319 <BRA>

A:Cross-references: EMBL:X17650; NID:G34829; PIDN:CNA35640.1; PID:G34830

C:Genetics:

A:Gene: GDB:MYOD1; MYF3

A:Cross-references: GDB:120707; OMIM:159970

A:Map position: 11p15.4-11p15.4

C:Superfamily: human myogenin

C:Keywords: DNA binding; transcription regulation

Query Match 100.0%; Score 87; DB 2; Length 319;

Best Local Similarity 100.0%; Pred. No. 3.9e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YSGPPSGARRNCYE 15

Db 212 YSGPPSGARRNCYE 226

RESULT 2

S20086

MyoD protein - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004

C:Accession: S20086

R:Huynen, L.; Bass, J.; Gardner, R.C.; Bellamy, A.R.

Nucleic Acids Res. 20, 374, 1992

A:Title: Nucleotide sequence of the sheep MyoD1 gene.

A:Reference number: S20086; MUID:92158635; PMID:1741269

A:Accession: S20086

Qy 1 YSGPPSGARRRNCYE 15
|||||:|||||:
Db 189 YSGPPCSSRRRNSYD 203

RESULT 7
A34783
myogenesis protein MyoDa - African clawed frog
N;Alternate names: Myoblast determination protein; myogenic factor 1
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Jul-1990 #sequence revision 13-Jul-1990 #text_change 09-Jul-2004
C;Accession: A34783; S06952; A60099
R;Scales, J.B.; Olson, E.N.; Perry, M.
Mol. Cell. Biol. 10, 1516-1524, 1990
A;Title: Two distinct Xenopus genes with homology to MyoD1 are expressed before somite f
A;Reference number: A34783; MUID:90205830; PMID:1690844
A;Accession: A34783
A;Molecule type: mRNA
A;Residues: 1-289 <SCA>
A;Cross-references: UNIPROT:P13904; GB:M31116
R;Hopwood, N.D.; Pluck, A.; Gurdon, J.B.
EMBO J. 8, 3409-3417, 1989
A;Title: MyoD expression in the forming somites is an early response to mesoderm induction
A;Reference number: S06952; MUID:90059936; PMID:2555164
A;Accession: S06952
A;Molecule type: mRNA
A;Residues: 1-289 <HOP>
A;Cross-references: EMBL:X16106; NID:G64906; PIDN:CAA34232.1; PID:G64907
R;Harvey, R.P.
Development 108, 669-680, 1990
A;Title: The Xenopus MyoD gene: an unlocalised maternal mRNA predates lineage-restricted
A;Reference number: A60099; MUID:90353188; PMID:2167198
A;Accession: A60099
A;Molecule type: mRNA
A;Residues: 1-289 <HAR>
C;Genetics:
A;Gene: mfl; MyoDa
C;Function:
A;Description: activates expression of skeletal muscle-specific genes
A;Pathway: myogenesis; skeletal muscle differentiation
A;Note: expressed exclusively in skeletal muscle and first activated during gastrulation
C;Superfamily: human myogenin
C;Keywords: alternative splicing; differentiation; DNA binding; skeletal muscle; transcr

Query Match 65.5%; Score 57; DB 2; Length 289;
Best Local Similarity 60.0%; Pred. No. 0.17;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 YSGPPSGARRRNCYE 15
|||||:|||||:
Db 184 YNSPPCSSRRRNSYD 198

RESULT 8
S15750
transforming protein (fra-1) - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S15750; I56863; S08010
R;Matsui, M.; Tokuhara, M.; Konuma, Y.; Nomura, N.; Ishizaki, R.
Oncogene 5, 249-255, 1990
A;Title: Isolation of human fos-related genes and their expression during monocyte-macro
A;Reference number: S15749; MUID:90191709; PMID:2107490
A;Accession: S15750
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-271 <MAT>
A;Cross-references: UNIPROT:P15407; EMBL:X16707; NID:G31462; PIDN:CAA34679.1; PID:G31463
R;Tsuchiya, H.; Fujii, M.; Niki, T.; Tokuhara, M.; Matsui, M.; Seiki, M.
J. Virol. 67, 7001-7007, 1993
A;Title: Human T-cell leukemia virus type 1 Tax activates transcription of the human fra
A;Reference number: I56863; MUID:94047311; PMID:8230424

A;Accession: I56863
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-5 <RES>
A;Cross-references: GB:D16365; NID:G537351; PIDN:BAA03867.1; PID:G4433154
C;Genetics:
A;Gene: fra-1
C;Superfamily: fos transforming protein; fos/jun DNA-binding domain homology
C;Keywords: DNA binding; nucleus
F;100-140/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 54.0%; Score 47; DB 2; Length 271;
Best Local Similarity 72.7%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GPPSGARRRNC 13
|||||:|||||:
Db 87 GPPPGVRRRPC 97

RESULT 9
TVRTFR
transforming protein fra-1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 09-Jul-2004
C;Accession: A27722; I57559
R;Cohen, D.R.; Curran, T.
Mol. Cell. Biol. 8, 2063-2069, 1988
A;Title: fra-1: a serum-inducible, cellular immediate-early gene that encodes a fos-rela
A;Reference number: A27722; MUID:88261282; PMID:3133553
A;Accession: A27722
A;Molecule type: mRNA
A;Residues: 1-275 <COH>
A;Cross-references: UNIPROT:P10158; GB:M19651; NID:G204174; PIDN:AAA41171.1; PID:G204175
A;Experimental source: Fischer F2408 embryo fibroblast cell line 208F
R;Bergers, G.; Graninger, P.; Bräselmann, S.; Wrighton, C.; Busslinger, M.
Mol. Cell. Biol. 15, 3748-3758, 1995
A;Title: Transcriptional activation of the fra-1 gene by AP-1 is mediated by regulatory
A;Reference number: I57559; MUID:95311973; PMID:7791782
A;Accession: I57559
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-43 <RES>
A;Cross-references: EMBL:U24154; NID:G1079522; PIDN:AAA82045.1; PID:G1079523
C;Comment: This protein is a fos-related antigen (fra) that may bind to DNA.
C;Genetics:
A;Gene: fra-1
A;Introns: 35/3
C;Superfamily: fos transforming protein; fos/jun DNA-binding domain homology
C;Keywords: DNA binding; leucine zipper; transforming protein
F;102-142/Domain: fos/jun DNA-binding domain homology <FJD>
F;135-163/Region: leucine zipper motif

Query Match 54.0%; Score 47; DB 1; Length 275;
Best Local Similarity 72.7%; Pred. No. 6;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GPPSGARRRNC 13
|||||:|||||:
Db 89 GPPPGVRRRPC 99

RESULT 10
H70761
probable dnapolymerase III - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: H70761
R;Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A,Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A,Reference number: A70500; MUID:98295987; PMID:9634230
A,Accession: H70761
A,Status: preliminary; nucleic acid sequence not shown; translation not shown
A,Molecule type: DNA
A,Molecule type: DNA
A,Residues: 1-1184 <COL>
A,Cross-references: UNIPROT:Q10779; GB:Z74020; GB:AL123456; NID:G261584; PIDN:CAA98315
A,Experimental source: strain H37Rv
C,Genetics:
C,Gene: dnaE1
C,Superfamily: DNA-directed DNA polymerase III alpha chain

Query Match 51.7%; Score 45; DB 2; Length 1184;
Best Local Similarity 57.1%; Pred. No. 41;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 SGPPSGARRNCYE 15
Db 335 AGPPDGYRERAAVE 348

```

RESULT 11
S11911
hypothetical protein - Streptomyces rimosus (fragment)
C/Species: Streptomyces rimosus
C/Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 09-Jul-2004
C/Accession: S11911
R/R: Plohl, M.; Gamulin, V.
Mol. Gen. Genet. 222, 129-134, 1990
A/Title: Five transfer RNA genes lacking CCA termini are clustered in the chromosome of
A/Reference number: S11911; MUID:91042413; PMID:1700267
A/Accession: S11911
A/Status: not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-127 <PLO>
A/Cross-references: UNIPROT:Q05400
A/Experimental source: strain R6-554

```

```
Query Match      50.6%; Score 44; DB 2; Length 127;
Best Local Similarity 88.9%; Pred. NO. 9.2;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

Qy 4 PPSGARRRN 12
Db 111 PPSGARRRS 119

RESULT 12

C34783 myogenesis protein MyoDb - African clawed frog
N:Alternate names: myogenic factor 25
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Jul-1990 #sequence revision 13-Jul-1990 #text_change 09-Jul-2004
R:Scales, J.B.; Olson, E.N.; Perry, M.
Mol. Cell. Biol. 10, 1516-1524, 1990
A:Title: Two distinct Xenopus genes with homology to MyoD1 are expressed before somitogenesis
A:Reference number: A34783; MUID:90205830; PMID:1690844
A:Accession: C34783
A:Molecule type: mRNA
A:Residues: 1-287 <SC>
A:Cross-references: UNIPROT:P16076; GB:M31118; NID:G214591; PIDN:AAA49902.1; PID:G214591
R:Harvey, R.P.
Development 108, 669-680, 1990
A:Title: The Xenopus MyoD gene: an unlocalised maternal mRNA predates lineage-restricted expression
A:Reference number: A60099; MUID:90353188; PMID:2167198
A:Accession: B60099
A:Molecule type: mRNA
A:Residues: 1-55, 'PE', 57-66, 'E', 67-98, 'E', 100-255, 'K', 257-281, 'N', 283-287 <HAR>
A:Note: authors translated the codon GAG for residue 101 as Arg
A:Accession: C60039
A:Status: not compared with conceptual translation
A:Molecule type: DNA

A;Residues: 2-55,'PE',57-182 <HA2>
C;Genetics:
A;Gene: mf25; Myodb
A;Introns: 182/3
A;Note: intron list may be incompl
C;Function:

A;Description: activates expression of skeletal muscle-specific genes
A;Pathway: myogenesis; skeletal muscle differentiation
A;Note: expressed exclusively in skeletal muscle and first activated during gastrulation
C;Superfamily: human myogenin
C;Keywords: differentiation; DNA binding; skeletal muscle

Query Match 50.6%; Score 44; DB 2; Length 287;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy	2	SGP	SGARR	NCYE	15
	:	:	:	:	:
Db	183	NSP	PCSSRR	NSYD	196

```

RESULT 13
T45957
hypothetical protein F7J8.150 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C/Accession: T45957
R/Beyan, M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Lem
submitted to the Protein Sequence Database, January 2000
A/Reference number: Z23018
A/Accession: T45957
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-368 <BEV>
A/Cross-references: UNIPROT:Q9LFE9; EMBL:AL137189
A/Experimental source: cultivar Columbia; BAC clone F7J8
C/Genetics:
A/Map position: 5
A/Note: F7J8.150

```

Query Match 50.6%; Score 44; DB 2; Length 568;
Best Local Similarity 58.3%; Pred. No. 32;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0

Qy 1 YSGPFGARRN 12
Dp 370 YTGPFSSRRKS 381

RESULT 14

E87371
hypothetical protein CC0985 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: E87371
R:Nierman, W.C.; Faldutylum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: E87371
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-949 <STO>
A:Cross-references: UNIPROT:Q9A9J5; GB:AE005673; MID:g13422269; PIDN:AAK22969.1; GSPDB:G
C:Genetics:
A:Gene: CC0985

Query Match 50.6%; Score 44; DB 2; Length 949;
Best Local Similarity 57.1%; Pred. No. 49;
Matches 8: Conservative 2; Mismatches 4: Indels

—

Qy 2 SGPPSGARRRNCYE 15
 ||||| ||| :
 Db 7 SGPPSFFRRRAHK 20

RESULT 15

T31322
 dCMP deaminase (EC 3.5.4.12) - Cenarchaeum symbiosum
 C:Species: Cenarchaeum symbiosum
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
 C:Accession: T31322
 E:Schleper, C.; DeLong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.
 J. Bacteriol. 180, 5003-5009, 1998
 A:Title: Genomic analysis reveals chromosomal variation in natural populations of the un
 A:Reference number: Z20994; MUID:98422450; PMID:9748430
 A:Accession: T31322
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-184 <SCH>
 A:Cross-references: UNIPROT:O74069; EMBL:AF083072; NID:G3599993; PID:G3599408; PIDN:AAC6
 C:Genetics:
 A:Gene: dcdt
 C:Keywords: hydrolase

Query Match 49.4%; Score 43; DB 2; Length 184;
 Best Local Similarity 53.3%; Pred. NO. 18;
 Matches 8; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

Qy 1 YSGPPSGARRRNCYE 15
 |:| | | :|||
 Db 54 YNGTPPGV--KNCYE 66

Search completed: March 3, 2005, 13:57:11
 Job time : 40 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 7, 2005, 03:44:32 ; Search time 1745 Seconds
(without alignments)
416.520 Million cell updates/sec

Title: US-10-018-964-3_COPY_1_15
Perfect score: 87
Sequence: 1 YSGPPSGARRNCYE 15

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_P/US10018964/runat_03032005_114706_21361/app_query.fasta_1.199
-DB=GenEmbl -QFMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10018964 @CGN 1.1 5600 @runat_03032005_114706_21361 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	87	100.0	963	9 BT007157	BT007157 Homo sapi
2	87	100.0	963	12 BT007461	BT007461 Synthetic
3	87	100.0	1416	9 HSMYF3	X17650 Human Myf-3
4	87	100.0	1692	9 HSMYOD	X56677 Human MyoD

5	87	100.0	1699	9	BC064493	Homo sapi
6	87	100.0	1757	6	CQ714840	Sequence
7	87	100.0	2480	6	AX677207	Sequence
8	87	100.0	3025	6	CQ787200	Sequence
9	87	100.0	3025	6	AX598705	Sequence
10	87	100.0	3025	6	AX767361	Sequence
11	87	100.0	3025	6	AX795665	Sequence
12	87	100.0	3025	6	AX822116	Sequence
13	87	100.0	3025	6	AX825756	Sequence
14	87	100.0	4086	6	I71382	Sequence 1
15	87	100.0	12825	9	AF027148	Homo sapi
16	87	100.0	115958	9	AC004736	Human Chr
17	87	100.0	115971	9	AC124056	Homo sapi
18	87	100.0	129199	2	AC005151	Homo sapi
19	87	100.0	145713	9	AC124301	Homo sapi
20	84	96.6	451	4	AF093675	Bos tauru
21	84	96.6	599	11	BV012581	myod1bts3
22	84	96.6	1055	4	AB110599	Bos tauru
23	84	96.6	2022	4	OAMYOD1	XG2102 O.aries mRN
24	84	96.6	2190	4	SSUI2574	U12574 Sus scrofa
25	64	73.6	3025	6	CQ787414	Sequence
26	64	73.6	3025	6	AX598843	Sequence
27	64	73.6	3025	6	AX767429	Sequence
28	64	73.6	3025	6	AX795747	Sequence
29	64	73.6	3025	6	AX822255	Sequence
30	64	73.6	3025	6	AX825895	Sequence
31	63	72.4	864	10	AF213003	Peromyscu
32	63	72.4	864	10	AF213004	Peromyscu
33	63	72.4	1785	10	MUSMYOD1A	M84918 M.musculus
34	63	72.4	1833	10	MUSMYOD	X61655 M.musculus
35	63	72.4	2627	10	MMYOD1	M84176 Rattus norv
36	63	72.4	6490	10	RATMYOD	AC026496 Homo sapi
37	63	72.4	120644	2	AC026496	AC103001 Rattus no
38	63	72.4	217732	2	AC103001	AC020786 Mus muscu
39	63	72.4	227330	10	AC020786	AC128786 Rattus no
40	63	72.4	234964	2	AC128786	AC132503 Rattus no
41	63	72.4	348540	2	AC132503	L16686 Coturnix co
42	60	69.0	1358	5	QULMF1A	AV641567 Meleagris
43	60	69.0	1435	5	AV641567	X16189 Chicken CMD
44	60	69.0	1518	5	GDCMD1	L34006 Chicken (cl
45	60	69.0	7389	5	CHKMYOD	

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
BT007157	Homo sapiens myogenic factor 3 mRNA, complete cds.	BT007157	BT007157.1	GI:30583152	FLJ_CDNA.	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	1 (bases 1 to 963)	Kalnina,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y., Phelan,M. and Farmer,A.	Cloning of human full-length CDSs in BD Creator(TM) System Donor vector	Unpublished	2 (bases 1 to 963)	Kalnina,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y., Phelan,M. and Farmer,A.	Direct Submission	Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA	This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two

forms: with and without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD In-Fusion(RM) cloning system between the SalI and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'ACC' after SalI site and before 'ATG', to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame. Clone distribution: <http://bioinfo.clontech.com/orclones>.

```

FEATURES             Location/Qualifiers
     1..963
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone_lib="GH00761X1.0"
         /clone_lib="BD Creator (TM) CDS Library derived from MGC
         collection"
         /lab_host="DH5alpha T1 resistant"
         /note="Vector: pDNR-Dual"
     1..963
         /codon_start=1
         /product="myogenic factor 3"
         /protein_id="AF35821.1"
         /db_xref="GI:30833153"
         /translation="MELLSPLRDVLTAPDGLCSFATTDVEYDDPCPSDPLRFFE
         DLPRLMHVGAALPKPEHSHFFPAHPVAPGAREDERVRAFGHQAGRCILWACKACK
         RKTNDRAAKAATMRRLNRYKNEAFETUKRTCTSSPNORLEKVELLRNAINIYIGL
         QALLRDQDAAPPGAAYAFAPGLPPGCGEHYSGSDASPSRNCSDGMMDVSGPPS
         GARRNCYEGAYNEAFPGPKGAASVSLDCLSSIVESTESPAAPALLADVPS
         ESPRRQEAAPASEGSGSGDPTQSPDAAPOCPAGANPNFIYQL"
CDS

```

```

ORIGIN

Alignment Scores:
Pred. No.:          0.0103          Length:          963
Score:             87.00           Matches:         15
Percent Similarity: 100.00%         Conservative:    0
Best Local Similarity: 100.00%       Mismatches:     0
Query Match:       100.00%          Indels:         0
DB:                9                Gaps:           0

US-10-018-964-3_COPY_1_15 (1-15) x BT007157 (1-963)

Qy      1  TyrSerGlyProProSerGlyAlaArgArgArgAncCysIYrGlu 15
Db      637  TACAGCGGCCCGCCGAGCGCGCGCCGCGCGGGGAACTGCTACGAA 681

```

RESULT 2	BT007461	963 bp	linear	SYN 13-MAY-2003
LOCUS	BT007461	963 bp	mRNA	
DEFINITION	Synthetic construct Homo sapiens myogenic factor 3 mRNA, partial cds.			
ACCESSION	BT007461			
VERSION	BT007461.1	GI:30583760		
KEYWORDS	FLI CDNA.			
SOURCE	synthetic construct			
ORGANISM	other sequences; artificial sequences.			
REFERENCE	1 (bases 1 to 963)			
AUTHORS	Kainine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y., Phelan,M. and Farmer,A.			
TITLE	Cloning of human full-length CDSs in BD Creator(TM) System Donor vector			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 963)			
AUTHORS	Kainine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y., Phelan,M. and Farmer,A.			
TITLE	Direct Submission			
JOURNAL	Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA			
COMMENT	This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal tags).			

tag). The CDS has been directionally cloned using BD In-Fusion (TM) cloning system between the SalI and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clones: 'ACC' after SalI site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame. Clone distribution: <http://bioinfo.clontech.com/orfclones>.

```

FEATURES
source
Location/Qualifiers
1..963
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/clone_lib="GH0076111.0"
/clone_lib="BD Creator (TM) CDS Library derived from MGC
collection"
/lab_host="DHSalpha T1 resistant"
/notes="vector: pDNR-Dual"
1..>963
/notes="Mutations: 962:Stop->Leu"
CDS
/transl_table=1
/codon_start=1
/product="Homo sapiens myogenic factor 3"
/protein_id="AAP36129.1"
/db_xref="GI:30583761"
/translation="MELLAPPLRDVLTAPDGLCSFATDDFVDDPCFSDPLRFEFE
DLDELPMVHGALUKPEEHGFFPAAPVAPFCAREDEVRAPSGHQAGRCILWACKACK
RKTNDARKAAATPERRRLSKYNEAPETLUKRTSSDNORLPKVELLRNAIRYIEGI
QALLLRQDAAAPGAAAFYAPGLPPFGGHEHYGSDSDSPRNCSDGMDWIDSGPFS
GARRNYCEGAYNEAPSPRPGKSAVSLDCLISIVERISTESPAPAPALLLADVPS
ESPSPROEAAAPGESGSDPTGSPDAAAPCCAGANPNPIYQVLL"

```

```

ORIGIN
Alignment Scores:
Pred. No.:          0.0103          Length:          963
Score:              87.00           Matches:         15
Percent Similarity: 100.00%         Conservative:    0
Best Local Similarity: 100.00%       Mismatches:     0
Query Match:        100.00%         Indels:         0
DB:                  12              Gaps:           0

US-10-018-964-3_COPY_1_15 (1-15) x BT007461 (1-963)

Qy      1  TyrSerGlyProProSerGlyAlaArgArgArgAsnCysTyrGlu 15
b6      517  TACAGCGCCCCCGAGCGCGCCCGCGCGCGGAACTGCTACGAA 681

```

RESULT 3	
HSMYF3	
LOCUS	1416 bp mRNA linear PRI 27-SEP-1990
DEFINITION	Human Myf-3 mRNA for myogenic determining factor 3'-fragment.
ACCESSION	HSMYF3
VERSION	X17650.1 GI:34829
KEYWORDS	developmental regulation; Myf gene; Myf-3 gene; MyoD1 protein; transcriptional activator.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1416) Braun, T., Bober E., Buschhausen-Denker, G., Kohtz, S., Grzeschik, K.H., Arnold, H.H. and Kotz, S.
TITLE	Differential expression of myogenic determination genes in muscle cells: possible autoactivation by the Myf gene products
JOURNAL	EMBO J. 8 (12), 3617-3625 (1989)
MEDLINE	90059960
PubMed	2583111
COMMENT	Data kindly reviewed (21-FEB-1990) by Harold H.H.
FEATURES	Location/Qualifiers
source	1..1416 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="1" /clone="Myf-3"

has 9777

CQ787200	LOCUS	CQ787200	3025 bp	DNA	linear	PAT 24-MAR-2004
	DEFINITION	Sequence 50 from Patent WO2004020662.				
	ACCESSION	CQ787200				
	VERSION	CQ787200.1	GI:45722183			


```
KEYWORDS      Homo sapiens (human)
SOURCE        Homo sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1
AUTHORS      Maier,S.
TITLE        Method and nucleic acids for the analysis of breast cell
              proliferative disorders
JOURNAL      Patent: WO 2004020662-A 50 11-MAR-2004;
              Epigenomics AG (DE)
FEATURES      source
              Location/Qualifiers
              1..3025
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:      0.0271      Length:      3025
Score:          87.00      Matches:      15
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             6          Gaps:          0
US-10-018-964-3_COPY_1_15 (1-15) x CQ787200 (1-3025)
Qy      1  TyrSerGlyProProSerGlyAlaArgArgAenCysTyrGlu 15
Db      1761 TACAGCGCGCCCGAGCGCGCGCGCGGAACTGCTACGAA 1805
RESULT 9
AX598705
LOCUS        AX598705      3025 bp      DNA      linear      PAT 14-FEB-2003
DEFINITION   Sequence 45 from Patent WO0207272.
ACCESSION    AX598705
VERSION      AX598705.1 GI:28398841
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1
AUTHORS      Berlin,K., Braun,A., Distler,J., Guetig,D., Howe,A., Mueller,J.,
              Olek,A., Piepenbrock,C., Adorjan,P., Grabs,G., Leache,R., Leu,E.,
              Lewin,A., Lipscher,E., Maier,S., Model,F., Mueller,V., Otto,T.,
              Pelet,C. and Ziebarth,H.
TITLE        Methods and nucleic acids for the analysis of hematopoietic cell
              proliferative disorders
JOURNAL      Patent: WO 0207272-A 45 03-OCT-2002;
              Epigenomics AG (DE)
FEATURES      source
              Location/Qualifiers
              1..3025
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:      0.0271      Length:      3025
Score:          87.00      Matches:      15
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             6          Gaps:          0
US-10-018-964-3_COPY_1_15 (1-15) x AX598705 (1-3025)
Qy      1  TyrSerGlyProProSerGlyAlaArgArgAenCysTyrGlu 15
Db      1761 TACAGCGCGCCCGAGCGCGCGCGCGGAACTGCTACGAA 1805
KEYWORDS      Homo sapiens (human)
SOURCE        Homo sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1
AUTHORS      Burger,M., Field,J.K., Genc,B., Liloglou,T., Lipscher,E., Maier,S.
              and Nimmrich,I.
TITLE        Method and nucleic acids for the analysis of a lung cell
              proliferative disorder
JOURNAL      Patent: WO 03052135-A 8 26-JUN-2003;
              Epigenomics AG (DE)
FEATURES      source
              Location/Qualifiers
              1..3025
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:      0.0271      Length:      3025
Score:          87.00      Matches:      15
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             6          Gaps:          0
US-10-018-964-3_COPY_1_15 (1-15) x AX795665 (1-3025)
Qy      1  TyrSerGlyProProSerGlyAlaArgArgAenCysTyrGlu 15
Db      1761 TACAGCGCGCCCGAGCGCGCGCGCGGAACTGCTACGAA 1805
RESULT 10
AX767361
LOCUS        AX767361      3025 bp      DNA      linear      PAT 02-JUL-2003
DEFINITION   Sequence 9 from Patent WO03044226.
ACCESSION    AX767361
VERSION      AX767361.1 GI:32435966
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1
AUTHORS      Burger,M., Caldwell,C., Genc,B., Becker,E., Maier,S. and
              Nimmrich,I.
TITLE        Method and nucleic acids for the analysis of a lymphoid cell
              proliferative disorder
JOURNAL      Patent: WO 03044226-A 9 30-MAY-2003;
              Epigenomics AG (DE)
FEATURES      source
              Location/Qualifiers
              1..3025
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:      0.0271      Length:      3025
Score:          87.00      Matches:      15
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             6          Gaps:          0
US-10-018-964-3_COPY_1_15 (1-15) x AX767361 (1-3025)
Qy      1  TyrSerGlyProProSerGlyAlaArgArgAenCysTyrGlu 15
Db      1761 TACAGCGCGCCCGAGCGCGCGCGCGGAACTGCTACGAA 1805
RESULT 11
AX795665
LOCUS        AX795665      3025 bp      DNA      linear      PAT 04-OCT-2003
DEFINITION   Sequence 8 from Patent WO03052135.
ACCESSION    AX795665
VERSION      AX795665.1 GI:37516331
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1
AUTHORS      Burger,M., Field,J.K., Genc,B., Liloglou,T., Lipscher,E., Maier,S.
              and Nimmrich,I.
TITLE        Method and nucleic acids for the analysis of a lung cell
              proliferative disorder
JOURNAL      Patent: WO 03052135-A 8 26-JUN-2003;
              Epigenomics AG (DE)
FEATURES      source
              Location/Qualifiers
              1..3025
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:      0.0271      Length:      3025
Score:          87.00      Matches:      15
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             6          Gaps:          0
US-10-018-964-3_COPY_1_15 (1-15) x AX795665 (1-3025)
```

QY 1 TySerGlyProProSerGlyAlaArgArgAsnCysTyrGlu 15
 Db 1761 TACAGCGGCCCCCGAGCGGCCCGCGCGGAACTGCTACGAA 1805

RESULT 12
 AX822116
 LOCUS AX822116 3025 bp DNA linear PAT 11-DEC-2003
 DEFINITION Sequence 8 from Patent EP1340818.
 ACCESSION AX822116
 VERSION AX822116.1 GI:39748744
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 Adorjan, P., Burger, M., Maier, S., Nimrich, I., Becker, E., Lesche, R.,
 Rujan, T. and Schmitt, A.
 TITLE Method and nucleic acids for the analysis of a colon cell
 proliferative disorder
 JOURNAL Patent: EP 1340818-A 8 03-SEP-2003;
 Epigenomics AG (DE)
 FEATURES Location/Qualifiers
 source
 1..3025
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 0.0271 Length: 3025
 Score: 87.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Db: 6 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x AX822116 (1-3025)

QY 1 TySerGlyProProSerGlyAlaArgArgAsnCysTyrGlu 15
 Db 1761 TACAGCGGCCCCCGAGCGGCCCGCGCGGAACTGCTACGAA 1805

RESULT 13
 AX825756
 LOCUS AX825756 3025 bp DNA linear PAT 11-DEC-2003
 DEFINITION Sequence 8 from Patent WO03072821.
 ACCESSION AX825756
 VERSION AX825756.1 GI:39751270
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 Adorjan, P., Burger, M., Maier, S., Nimrich, I., Becker, E., Lesche, R.,
 Rujan, T. and Schmitt, A.
 TITLE Method and nucleic acids for the analysis of a colon cell
 proliferative disorder
 JOURNAL Patent: WO 03072821-A 8 04-SEP-2003;
 Epigenomics AG (DE)
 FEATURES Location/Qualifiers
 source
 1..3025
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 0.0271 Length: 3025
 Score: 87.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Db: 6 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x AX825756 (1-3025)

QY 1 TySerGlyProProSerGlyAlaArgArgAsnCysTyrGlu 15
 Db 1761 TACAGCGGCCCCCGAGCGGCCCGCGCGGAACTGCTACGAA 1805

RESULT 14
 AX825756
 LOCUS AX825756 3025 bp DNA linear PAT 03-APR-1998
 DEFINITION Sequence 1 from patent US 5681735.
 ACCESSION I71382
 VERSION I71382.1 GI:3007517
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 4086)
 AUTHORS Emerson, C.P. and Goldhamer, D.J.
 TITLE Transcription control element for increasing gene expression in
 myoblasts
 JOURNAL Patent: US 5681735-A 1 28-OCT-1997;
 FEATURES Location/Qualifiers
 source
 1..4086
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Alignment Scores:
 Pred. No.: 0.0349 Length: 4086
 Score: 87.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Db: 6 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x I71382 (1-4086)

QY 1 TySerGlyProProSerGlyAlaArgArgAsnCysTyrGlu 15
 Db 2399 TACAGCGGCCCCCGAGCGGCCCGCGCGGAACTGCTACGAA 2443

RESULT 15
 AF027148
 LOCUS AF027148 12825 bp DNA linear PRI 07-AUG-1998
 DEFINITION Homo sapiens myogenic determining factor 3 (MYOD1) gene, complete
 cds.
 ACCESSION AF027148
 VERSION AF027148.1 GI:3403164
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 12825)
 AUTHORS Chen, B., Dias, P., Jenkins, J. J., Savell, V. H. and Parham, D. M.
 TITLE Methylation alterations of the MyoD1 upstream region are predictive
 of subclassification of human rhabdomyosarcomas
 JOURNAL Am. J. Pathol. 152 (4), 1071-1079 (1998)
 MEDLINE 98206444
 PUBMED 9546368
 REFERENCE 2 (bases 1 to 12825)
 AUTHORS Chen, B.
 TITLE Direct Submission
 JOURNAL Submitted (26-SEP-1997) Pathology, University of Arkansas for
 Medical Sciences, 4301 West Markham St., Little Rock, AR 72205, USA
 COMMENT Methylation alterations in the 5' region are found in embryonal
 rhabdomyosarcoma and alveolar rhabdomyosarcoma. Dynamic
 methylation changes in this region are found in normal myogenesis.
 FEATURES Location/Qualifiers

Handwritten notes: "X-axis: myoblasts" and "BMT-H" (with an arrow pointing to the AF027148 entry).

source 1. .12825
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11p15.4"
repeat_region 1866. .2513
/note="VNTR; polymorphic in the general population"
/rpt_type=tandem
8721. .12783
gene /gene="MYOD1"
/note="expressed in early myogenesis and in
rhabdomyosarcomas"
mRNA join(10264. .11065,11555. .11633,11908. .12783)
/gene="MYOD1"
/product="myogenic determining factor 3"
CDS join(10436. .11065,11555. .11633,11908. .12161)
/gene="MYOD1"
/note="Myf3; MyoD1"
/codon_start=1
/product="myogenic determining factor 3"
/protein_id="AAC29001.1"
/db_xref="GI:3403165"
/translation="MELLSPLRDVLTAPDGLCSFATTTDDFYDDFCFDSPLRFEE
DLDPRLMHVGALLKPEESHSHFPAAVHPAGAREDEHVRAPSGHHGKCLLWACKCK
RKTNADRRKAATNRERRLSKVNEAFETLKRCTSSNPQRLPKVEILRNAYRIEGL
QALLRDQAAAPPAGAAAFYAPGLPPGRGGEHYSGDSASPRSCSDGMWDYSGPPS
GARRNCVEGAYNEAPSEPPGKSAVSLDCLSSIVRISTESPAAPALLADVPS
ESPPRQEAAPSESGSDPTQSPDAAPQCPAGANPNPIYQVL"

ORIGIN

Alignment Scores:
Pred. No.: 0.0915 Length: 12825
Score: 87.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x AF027148 (1-12825)

Qy 1 TyrSerGlyProProSerGlyAlaArgArgArgAsnCysTyrGlu 15
|||||
Db 11561 TACAGCGCCCCCGAGCGCGCGCGCGGGAACCTGCTACGAA 11605
|||||

Search completed: March 7, 2005, 09:09:09
Job time : 1751 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 6, 2005, 23:13:02 ; Search time 427 Seconds
(without alignments)
207.953 Million cell updates/sec

Title: US-10-018-964-3_COPY_1_15

Perfect score: 87
Sequence: 1 YSGPPSGARRRNCYE 15

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10018964/runat_03032005_114706_21351/app_query.fasta_1.199
-DB=N_Geneseq_16Dec04 -QMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10018964 @CGN 1.1 708 @runat_03032005_114706_21351 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=110 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	100.0	1757	13 ADP07312	Adp07312 Human MYO
2	87	100.0	1757	12 ACN39187	Actn39187 Tumour-as
3	87	100.0	2480	8 ADA84118	Ada84118 Human MYO
4	87	100.0	2744	12 ADQ25112	Adq25112 Human sof
5	87	100.0	3025	8 ABZ09905	Abz09905 Human 5'

6	87	100.0	3025	10 ADB53952	Adb53952 MYOD1 gen
7	87	100.0	3025	10 ADE84013	Ade84013 5' regula
8	87	100.0	4086	2 AAG51033	Aag51033 Human myo
9	64	73.6	3025	8 ABZ10043	Abz10043 Haematopo
10	64	73.6	3025	10 ADB54091	Adb54091 Pretreat
11	64	73.6	3025	10 ADE84081	Ade84081 Human lym
12	63	72.4	957	3 AAA47075	Aaa47075 cDNA sequ
13	63	72.4	1849	8 ACC47992	Acc47992 Mouse Myo
14	63	72.4	3636	2 AAQ74447	Aaq74447 Proximal
15	56	64.4	3025	8 ABZ10044	Abz10044 Haematopo
16	56	64.4	3025	10 ADB54092	Adb54092 Pretreat
17	56	64.4	3025	10 ADE84082	Ade84082 Human lym
18	53	60.9	1097	8 ABZ52395	Abz52395 Aspergill
19	52	59.8	432	11 ABD03483	Abd03483 Pseudomon
20	52	59.8	542	5 ABV12557	Abv12557 Human pro
21	52	59.8	987	11 ABD03199	Abd03199 Pseudomon
22	52	59.8	1272	11 ABD03272	Abd03272 Pseudomon
23	52	59.8	2358	8 ACA25942	Acta25942 Prokaryot
24	51	58.6	303	11 ABD00514	Abd00514 Klebsiell
25	51	58.6	603	9 ADB09169	Adb09169 Alloiococ
26	51	58.6	603	9 ADB09171	Adb09171 Alloiococ
27	51	58.6	855	8 ABT20902	Abt20902 Aspergill
28	51	58.6	855	8 ABT15082	Abt15082 Aspergill
29	51	58.6	908	8 ABT20304	Abt20304 Aspergill
30	51	58.6	908	8 ABT18488	Abt18488 Aspergill
31	51	58.6	1233	12 ADM80109	Adm80109 Spiramyci
32	51	58.6	1233	12 ADN97625	Adn97625 S ambofac
33	51	58.6	1865	5 AAS71990	Aas71990 DNA encod
34	51	58.6	2908	8 ABT19708	Abt19708 Aspergill
35	51	58.6	2908	8 ABT17894	Abt17894 Aspergill
36	51	58.6	30943	12 ADM80034	Adm80034 Spiramyci
37	51	58.6	30943	12 ADN97550	Adn97550 S ambofac
38	51	58.6	110000	9 ADB12064_08	Continuation (9 of
39	50	57.5	202	10 ACA55687	Acta55687 Rabbit si
40	50	57.5	202	12 ADI55483	Adi55483 Human pol
41	50	57.5	286	10 ABZ90108	Abz90108 Human nuc
42	50	57.5	286	11 ABD26338	Abd26338 AA459692
43	50	57.5	1568	6 ABA92359	Abag92359 Human glu
44	50	57.5	2691	12 ADP99229	Adp99229 Human tra
45	50	57.5	2977	4 AAD17478	Aad17478 Human tra

ALIGNMENTS

RESULT 1

ADP07312

ID ADP07312 standard; DNA; 1757 BP.

XX AC ADP07312;

XX AC

DT 29-JUL-2004 (first entry)

XX XX

DE Human MYOD1 DNA.

XX XX

KW ds; proliferative disease; breast; methylation; CpG; bisulfite; human.

XX OS Homo sapiens.

XX XX

FN DE10255104-Al.

XX XX

PD 11-MAR-2004.

XX XX

PF 26-NOV-2002; 2002DE-01055104.

XX XX

PR 27-AUG-2002; 2002DE-01039313.

XX XX

PA (EPIC-) EPIGENOMICS AG.

XX XX

PI Maier S;

XX XX

DR WPI; 2004-284340/27.

XX GENBANK; NM_002478.

OS Homo sapiens.
 XX WO200277272-A2.
 XX 03-OCT-2002.
 XX 26-MAR-2002; 2002WO-EP003401.
 XX 26-MAR-2001; 2001US-0278333P.
 XX (EPIC-) EPIGENOMICS AG.
 XX Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
 PI Olek A, Pipebrock C, Adorjan P, Grabs G, Lesche R, Leu E;
 PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelat C;
 PI Schwoppe I, Ziebarth H;
 XX WPI; 2003-018942/01.
 XX Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent that
 PT distinguishes between methylated and non-methylated CpG dinucleotides.
 XX Claim 38; SEQ ID NO 45; 117pp; English.
 XX The present invention describes a method for detecting and
 CC differentiating between hematopoietic cell proliferative disorders
 CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used: for
 CC differentiating between healthy hematopoietic cells and proliferative
 CC disorder hematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of hematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of hematopoietic cell proliferation disorder related DNA
 CC sequences. The nucleotide sequences from the present invention can also
 CC be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC hematopoietic cell proliferative disorders. The present method enables a
 CC highly specific classification of hematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients
 XX
 SQ Sequence 3025 BP; 535 A; 1012 C; 930 G; 548 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0.0118 Length: 3025
 Score: 87.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 8 Gaps: 0
 US-10-018-964-3_COPY_1_15 (1-15) x ABZ09905 (1-3025)
 QY 1 TyrSerGlyProProSerGlyAlaArgArgArgArgCysTyrGlu 15
 DB 1761 TACAGCGCGCCCGAGCGCGCCCGCGCGGAACTGCTACGAA 1805
 RESULT 6
 ADB53952
 ID ADB53952 standard; DNA; 3025 BP.
 XX
 AC ADB53952;
 DT 04-DEC-2003 (first entry)
 XX

DE MYOD1 genomic DNA region.
 XX
 KW colon cell proliferative disorder; non methylated CpG dinucleotide;
 KW cytosatic; cancer; adenoma; carcinoma; cytosine methylation state; ds.
 XX
 OS Unidentified.
 XX
 PN WO2003072821-A2.
 XX
 PD 04-SEP-2003.
 XX
 XX 27-FEB-2003; 2003WO-EP002035.
 XX
 PR 27-FEB-2002; 2002EP-00004551.
 XX
 PA (EPIC-) EPIGENOMICS AG.
 XX
 XX Adorjan P, Burger M, Maier S, Nimmrich I, Becker E, Lesche R;
 PI Rujan T, Schmitt A;
 PI
 XX WPI; 2003-731620/69.
 DR
 XX Detecting and differentiating between colon cell proliferative disorders
 PT associated with a gene or its regulatory regions comprises contacting a
 PT target nucleic acid in a biological sample obtained from the subject with
 PT a reagent.
 XX
 PS Claim 46; SEQ ID NO 8; 74pp; English.
 XX
 CC The invention relates to a novel method for detecting and differentiating
 CC between colon cell proliferative disorders associated with at least one
 CC gene or its regulatory regions. The method comprises contacting a target
 CC nucleic acid in a biological sample obtained from the subject with at
 CC least one reagent or a series of reagents, where the reagent or series of
 CC reagents, distinguishes between methylated and non methylated CpG
 CC dinucleotides within the target nucleic acid. The molecules of the
 CC invention demonstrate cytostatic activity whilst the method may useful
 CC for detecting and differentiating between colon cell proliferative
 CC disorders, including cancers such as colon adenoma and colon carcinoma.
 CC The PNA (peptide nucleic acid)-oligomers are useful as probes for
 CC determining cytosine methylation state or single nucleotide
 CC polymorphisms. The current sequence is that of the genomic DNA region of
 CC the invention. This sequence is not shown within the specification but is
 CC taken from Wipoweb.
 XX
 SQ Sequence 3025 BP; 535 A; 1012 C; 930 G; 548 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0.0118 Length: 3025
 Score: 87.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0
 US-10-018-964-3_COPY_1_15 (1-15) x ADB53952 (1-3025)
 QY 1 TyrSerGlyProProSerGlyAlaArgArgArgArgCysTyrGlu 15
 DB 1761 TACAGCGCGCCCGAGCGCGCCCGCGCGGAACTGCTACGAA 1805
 RESULT 7
 ADE84013
 ID ADE84013 standard; DNA; 3025 BP.
 XX
 AC ADE84013;
 DT 29-JAN-2004 (first entry)
 XX
 DE 5' regulatory region of human gene MYOD1.
 XX
 KW ds; lymphoid cell proliferative disorder; methylation;
 KW methylated CpG dinucleotide; single nucleotide polymorphism; SNP;

diffuse large B-cell lymphoma; mantle cell lymphoma;
chronic lymphocytic leukemia; small lymphocytic lymphoma;
follicular lymphoma; diagnosis; prognosis.

Homo sapiens.

WO2000304226-A2.

30-MAY-2003.

25-NOV-2002; 2002WO-EP013265.

23-NOV-2001; 2001DE-01057491.

28-DEC-2001; 2001DE-01064501.

(EPIG-) EPIGENOMICS AG.

Burger M, Caldwell C, Genc B, Becker E, Maier S, Nimmrich I;

WPI; 2003-457621/43.

Detecting and differentiating between lymphoid cell proliferative disorders comprises contacting a target nucleic acid with at least one reagent that distinguishes between methylated and non-methylated CpG dinucleotides.

Disclosure; SEQ ID NO 9; 448pp; English.

The invention relates to a method of detecting and differentiating between lymphoid cell proliferative disorders associated with at least one gene and/or their regulatory regions in a subject by contacting a target nucleic acid in a biological sample obtained from the subject with at least one reagent or series of reagents that distinguish between methylated and non-methylated CpG dinucleotides within the target nucleic acid. The genes and/or their regulatory regions are preferably selected from MDRI, CSNK2B, EGR4, AR, CDK4, RB2, CDC25A, GP1B beta, MYOD1, CDH3, MYCL1, ELK1, ABL1, APC, BCL2, CDH1, CDKN1A, CDKN1B, CDKN2A, FOS, GSTP1, HIC-1, MGMT, MLH1, MOS, MYC, PTEN, RBL2, TGFBR2, TP73, CDKN1C, GSK3beta, ESRI, APAF1, BAK1, BAX or HOXA5. Oligomers, peptide nucleic acid (PNA)-oligomers and/or isolated nucleic acids based on the sequences of the genes are useful for detecting the methylation state of all the CpG dinucleotides within one or more the sequences, or their complements, for determining the cytosine methylation state and or single nucleotide polymorphisms (SNPs), and for differentiating at least two of the medical conditions such as diffuse large B-cell lymphoma, mantle cell lymphoma, chronic lymphocytic leukemia, small lymphocytic lymphoma and follicular lymphoma. They are also useful for detecting of a predisposition to, differentiation between subclasses, diagnosis, prognosis, treating and/or monitoring of lymphoid cell proliferative disorder. This sequence represents the 5' and/or regulatory region from one of the above mentioned genes.

Sequence 3025 BP; 535 A; 1012 C; 930 G; 548 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.0118 Length: 3025
Score: 87.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x ADB84013 (1-3025)

Qy 1 TyrSerGlyProProSerGlyAlaArgArgArgenCysTyrGlu 15
Db 1761 TACAGCGCGCCCGAGCGCGCGCGCGGCGGAACTGCTACGAA 1805

RESULT 8

AAQ51033

ID AAQ51033 standard; DNA; 4086 BP.

XX

AC AAQ51033;

XX

DT

DT

XX

DE

XX

KW

KW

XX

OS

XX

XX

Key

TATA_signal

FT

FT

FT

FT

FT

FT

FT

PN

XX

XX

PD

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

25-MAR-2003 (revised)

09-MAY-1994 (first entry)

Human myoD gene and myoblast specific transcriptional control element.

Enhancer; myogenic cells; expression; 5'; control element;

antisense oligonucleotides; gene therapy; ss.

Homo sapiens.

Key

Location/Qualifiers

1038..1044

/*tag= a

promoter

1240

/*tag= b

/note= "3' end of promoter region"

1277..1652

/*tag= c

CDS

W09321347-A1.

28-OCT-1993.

24-MAR-1993; 93WO-US002767.

10-APR-1992; 92US-00866386.

(FOX-) FOX CHASE CANCER CENT.

Emerson CP, Goldhamer DJ;

WPI; 1993-351756/44.

P-PSDB; AAR42359.

Transcription control element with enhancer activity - increases gene expression in myoblast(s) and is derived from upstream region of myoD gene.

Claim 1; Page 16-18; 56pp; English.

The sequence is that of the human myoD gene and its upstream regulatory sequence. The transcriptional control element has enhancer activity in myogenic cells, causing increased expression of the myoD gene. Antisense oligonucleotide fragments of the control sequence can be used for identifying and locating partic. functional regions in the control element. The myoblast-specific control element will aid advances in gene therapy using myoblast transfer and microinjection techniques. The control element may also be useful in biochemical assays for the activity of early transcription factors involved in regulating enhancers. See also AAQ51034-5. (Updated on 25-MAR-2003 to correct PN field.)

Sequence 4086 BP; 769 A; 1306 C; 1206 G; 805 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.0163 Length: 4086

Score: 87.00 Matches: 15

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 2 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x AAQ51033 (1-4086)

Qy 1 TyrSerGlyProProSerGlyAlaArgArgArgenCysTyrGlu 15

Db 2399 TACAGCGCGCCCGAGCGCGCGCGGCGGAACTGCTACGAA 2443

RESULT 9

ABZ10043

ID ABZ10043 standard; DNA; 3025 BP.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

AD84081;
29-JAN-2004 (first entry)
Human lymphoid cell proliferative disorder gene derived DNA #17.
ds; lymphoid cell proliferative disorder; methylation;
methylated CpG dinucleotide; single nucleotide polymorphism; SNP;
diffuse large B-cell lymphoma; mantle cell lymphoma;
chronic lymphocytic leukemia; small lymphocytic lymphoma;
follicular lymphoma; diagnosis; prognosis.
Homo sapiens.
WO2003044226-A2.
30-MAY-2003.
25-NOV-2002; 2002WO-EP013265.
23-NOV-2001; 2001DE-01057491.
28-DEC-2001; 2001DE-01064501.
(EPiG-) EPIGENOMICS AG.
Burger M, Caldwell C, Genc B, Becker E, Maier S, Nimmrich I;
WPI; 2003-457621/43.
Detecting and differentiating between lymphoid cell proliferative
disorders comprises contacting a target nucleic acid with at least one
reagent that distinguishes between methylated and non-methylated CpG
dinucleotides.
Claim 26; SEQ ID NO 77; 448pp; English.
The invention relates to a method of detecting and differentiating
between lymphoid cell proliferative disorders associated with at least
one gene and/or their regulatory regions in a subject by contacting a
target nucleic acid in a biological sample obtained from the subject with
at least one reagent or series of reagents that distinguish between
methylated and non-methylated CpG dinucleotides within the target nucleic
acid. The genes and/or their regulatory regions are preferably selected
from MDR1, CSNK2B, EGR4, AR, CDK4, RB2, CDC25A, GP1b beta, MYO01, CDH3,
MYCL1, ELK1, ABL1, APC, BCL2, CDH1, CDKN1B, CDKN2a, CDKN2B, FOS,
GSTP1, HIC-1, MGMT, MLH1, MOS, MYC, PTEN, RBL2, TGFBR2, TP73, CDKN1C,
GSK3beta, ESR1, APAF1, BAK1, BAX or HOXA5. Oligomers, peptide nucleic
acid (PNA)-oligomers and/or isolated nucleic acids based on the sequences
of the genes are useful for detecting the methylation state of all the
CpG dinucleotides within one or more the sequences, or their complements,
for determining the cytosine methylation state and/or single nucleotide
polymorphisms (SNPs), and for differentiating at least two of the medical
conditions such as diffuse large B-cell lymphoma, mantle cell lymphoma,
chronic lymphocytic leukemia, small lymphocytic lymphoma and follicular
lymphoma. They are also useful for detecting of a predisposition to,
differentiation between subclasses, diagnosis, prognosis, treating and/or
monitoring of lymphoid cell proliferative disorder. This sequence
represents a nucleic acid of a pretreated genomic DNA derived from the
above mentioned genes.
SQ Sequence 3025 BP; 535 A; 245 C; 930 G; 1315 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 24-5 Length: 3025
Score: 64.00 Matches: 12
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 3
Query Match: 73-56% Indels: 0
DB: 10 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x AD84081 (1-3025)
Qy 1 TyrserGlyProProSerGlyAlaArgGArgAsnCysTyrGlu 15

Db 1761 TATAGCGGTTTTCGAGCGGCGTTTCGGCGGGAATTGTTACGAA 1805
RESULT 12
AAA47075
ID AAA47075 standard; cDNA; 957 BP.
XX
AC AAA47075;
XX
XX 03-OCT-2000 (first entry)
XX cDNA sequence encoding the transcription factor protein MyoD1.
DE Transcription factor; MyoD1; myoD1 gene; regulatory gene;
KW skeletal myogenesis; DNA-binding protein; bHLH domain;
KW augmentor sequence; ss.
XX
OS Mus sp.
XX
XX Key Location/Qualifiers
FH 1..957
CDS /*tag= a
FT /transl_except= (pos: 385..387, aa: Pro)
FT /transl_except= (pos: 415..417, aa: Asp)
FT /transl_except= (pos: 421..423, aa: Asp)
FT /transl_except= (pos: 454..456, aa: Asp)
FT /transl_except= (pos: 667..669, aa: Asp)
FT /product= "transcription factor protein MyoD1"
XX
XX WO200037682-A1.
PN
XX
XX 29-JUN-2000.
XX
XX 20-DEC-1999; 99WO-US030460.
XX
XX 21-DEC-1998; 98US-0113261P.
XX
XX (PROM-) PROMEGA CORP.
PA (SCHE/) SCHENBORN E T.
PA (BRON/) BRONDYK W.
PA (DBER/) DE BERG L G.
PA (ALMO/) ALMOND B D.
XX
XX Schenborn ET, Brondyk W, De Berg LG, Almond BD;
XX
XX WPI; 2000-442694/38.
DR P-PSDB; AAY93769.
XX
XX Improving transient expression of an extrachromosomal foreign gene in
PT cultured eukaryotic cells to produce large amounts of a transgenic gene
XX product in cultured cells.
XX
XX Disclosure; Fig 1A; 133pp; English.
XX
XX The present sequence encodes a murine the transcription factor protein
CC MyoD1. It is derived from the myoD1 gene, which is a regulatory gene for
CC skeletal myogenesis. The protein is a DNA-binding protein that binds to
CC the enhancer elements of several muscle-specific genes. The myoD1 gene
CC contains a bHLH domain, which consists of a basic segment that binds to a
CC consensus DNA sequence (CANNTG) element termed the "E-box", and a short
CC alpha helix connected by a loop to a longer alpha helix. The loop is
CC flexible and allows the two helices to fold back against each other. The
CC alpha helix binds to DNA and the HLH structure of another bHLH
CC polypeptide. The second polypeptide can be the same or different, and the
CC resulting protein complexes then bind to DNA in the 5' regulatory regions
CC of genes. An augmentor sequence can be derived from the myoD1 cDNA
CC sequence, which augments post-transfection expression of a foreign gene
CC of interest. The augmentor sequence is used in the method of the
CC invention, for the improving transient gene expression of an
CC extrachromosomal foreign gene in a host cell
XX
XX Sequence 957 BP; 182 A; 341 C; 282 G; 152 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9.92 Length: 957
Score: 63.00 Matches: 11
Percent Similarity: 86.67% Conservative: 2
Best Local Similarity: 73.33% Mismatches: 2
Query Match: 72.41% Indels: 0
DB: 3 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x AAA47075 (1-957)

QY 1 TyfSerGlyProProSerGlyAlaArgArgArgAsnCysTyrGlu 15
|||||
Db 634 TACAGCGGCCCCCAAGCGGCCCCCGCGGCGAGATGGCTACGAC 678
|||||

RESULT 13
ACCA47992
ID ACCA47992 standard; DNA; 1849 BP.
XX ACCA47992;
AC ACCA47992;
DT 11-AUG-2003 (first entry)
XX
DE Mouse MyoD nucleotide sequence.
XX
KW Cell differentiation; gene expression; neuroprotective; immunomodulator;
KW dermatological; nontropic; antiparkinsonian; antianemic; cytostatic;
KW anti-HIV; protozoicide; vulnery; deacetylase; MyoD; gene; ds.
XX
OS Mus musculus.
XX
PN W02003033678-A2.
XX
PD 24-APR-2003.
XX
PF 17-OCT-2002; 2002WO-US033570.
XX
PR 18-OCT-2001; 2001US-0335705P.
PR 25-OCT-2001; 2001US-0343854P.
XX
XX (SALK) SALK INST BIOLOGICAL STUDIES.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Sartorelli V, Puri PL;
PI WPI; 2003-430347/40.
XX
DR
XX
PT Enhancing progenitor cell differentiation and regeneration or
PT differentiation-related gene expression in a progenitor cell, useful for
PT treating tissue degeneration, comprises contacting the cell with a
PT deacetylase inhibitor.
XX
PS Disclosure; Page 69-70; 79pp; English.

XX The invention relates to enhancing progenitor cell differentiation or
XX differentiation-related gene expression in a progenitor cell. The method
XX involves contacting an undifferentiated progenitor cell with an amount of
XX a deacetylase inhibitor for a period of time sufficient to induce
XX progenitor cell differentiation or enhance expression of the genes. The
XX method is useful in promoting cell differentiation and regeneration using
XX deacetylase inhibitors. The method is used to inhibit, prevent or treat
XX diseases or conditions associated with a degeneration or loss of tissue,
XX such as muscle tissue, nerve tissue or haematopoietic tissue. In
XX particular, the disease or condition is muscular atrophy, muscular
XX dystrophy, muscular cachexia, dermatomyositis, Alzheimer's disease,
XX olivopontocerebellar atrophy, Parkinson's disease, degeneration of
XX nervous tissue, ocular atrophy, hepatocerebral degeneration, idiopathic
XX aplastic anemia, secondary aplastic anemia, amyotrophic lateral
XX sclerosis, poliomyelitis, bone marrow loss induced by radiation therapy
XX or chemotherapy, multiple myeloma, acute lymphocytic leukemia, HIV
XX infection, AIDS, malaria, chronic myelogenous leukemia, Fanconi's anemia
XX or trauma. The present sequence represents a mouse MyoD DNA (GenBank
XX Accession No. XM124916)

SQ Sequence 1849 BP; 376 A; 584 C; 506 G; 393 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 20.1 Length: 1849
Score: 63.00 Matches: 11
Percent Similarity: 86.67% Conservative: 2
Best Local Similarity: 73.33% Mismatches: 2
Query Match: 72.41% Indels: 0
DB: 8 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x ACCA47992 (1-1849)

QY 1 TyfSerGlyProProSerGlyAlaArgArgArgAsnCysTyrGlu 15
|||||
Db 825 TACAGCGGCCCCCAAGCGGCCCCCGCGGCGAGATGGCTACGAC 869
|||||

RESULT 14
AAQ74447
ID AAQ74447 standard; DNA; 3636 BP.
XX AAQ74447;
AC AAQ74447;
XX
DT 19-MAY-1995 (first entry)
XX
DE Proximal regulatory region of the MyoD gene.
XX
KW MyoD; muscle; gene expression; myoblast; muscle differentiation;
KW rhabdomyoma cell; nuclear regulatory factor; chemotherapeutic agent;
KW identification; detection; skeletal muscle; ss.
XX
OS Mus musculus.
XX
PN US5352595-A.
XX
PD 04-OCT-1994.
XX
PF 03-SEP-1991; 91US-00753520.
XX
PR 03-SEP-1991; 91US-00753520.
XX
PA (HUTC-) HUTCHINSON CANCER RES CENT FRED.
XX
PI Palmer TD, Tapscott SJ, Weintraub HM;
XX WPI; 1994-316186/39.
XX
DR
XX
PT Isolated proximal and distal myoD regulatory regions - and vectors contg.
PT then, for tissue selective gene expression in muscle cells and for
PT positive or negative muscle cell selection.
XX
PS Claim 1; Fig 1; 48pp; English.

XX The myoD proximal and distal regulatory regions (See AAQ74447, AAQ74448)
XX may be inserted into vectors and used (1) to induce a muscle phenotype in
XX a non-muscle cell, (2) for positive or negative selection of muscle cells
XX (3) for targeted gene expression, specifically in skeletal muscle, (4)
XX for the expression of genes controlling growth of myoblasts or mutant
XX forms of myoD that modulate muscle differentiation, (5) for engineering
XX rhabdomyoma cells to increase sensitivity to chemotherapeutic agents,
XX (6) for the expression of growth factors to stimulate skeletal muscle
XX growth in animals, and (7) to identify nuclear regulatory factors that
XX bind to the regulatory region

SQ Sequence 3636 BP; 730 A; 1096 C; 990 G; 814 T; 0 U; 6 Other;

Alignment Scores:
Pred. No.: 41.6 Length: 3636
Score: 63.00 Matches: 11
Percent Similarity: 86.67% Conservative: 2
Best Local Similarity: 73.33% Mismatches: 2
Query Match: 72.41% Indels: 0
DB: 2 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x AAQ74447 (1-3636)
Qy 1 TyrSerGlyProProSerGlyAlaArgArgArgAenCysTyrGlu 15
Db 2133 TACAGCGGCCCCCAAGCGCCCCCGCGCGGAGATGCTACGAC 2177

RESULT 15

ABZ10044/c
ID ABZ10044 standard; DNA; 3025 BP.
XX
AC ABZ10044;
XX
DT 16-JAN-2003 (first entry)
XX
DE Haematopoietic cell proliferation disorder related DNA sequence #184.
XX
KW Human; haematopoietic cell proliferation disorder; cytostatic;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200277272-A2.
XX
PD 03-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-EP003401.
XX
PR 26-MAR-2001; 2001US-0279333P.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI Schwöpe I, Ziebarth H;
XX
DR WPI; 2003-018942/01.
XX

PT Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated CpG dinucleotides.

PS Claim 28; SEQ ID NO 184; 117pp; English.

CC The present invention describes a method for detecting and
CC differentiating between haematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used: for
CC differentiating between healthy haematopoietic cells and proliferative
CC disorder haematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC determining the cytosine methylation state and/or single nucleotide
CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of haematopoietic cell proliferation disorder related
CC sequences. The nucleotide sequences from the present invention can also
CC be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC haematopoietic cell proliferation disorders. The present method enables a
CC highly specific classification of haematopoietic cell proliferative
CC disorders allowing for improved and informed treatment of patients

SQ Sequence 3025 BP; 548 A; 245 C; 1012 G; 1220 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 349 Length: 3025
Score: 56.00 Matches: 10

Percent Similarity: 80.00% Conservative: 2
Best Local Similarity: 66.67% Mismatches: 3
Query Match: 64.37% Indels: 0
DB: 8 Gaps: 0
US-10-018-964-3_COPY_1_15 (1-15) x ABZ10044 (1-3025)
Qy 1 TyrSerGlyProProSerGlyAlaArgArgArgAenCysTyrGlu 15
Db 1365 TACAACGACCCCGGACGCGCGCGGAGAACTACTACGAA 1221

Search completed: March 7, 2005, 08:39:51
Job time : 433 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 7, 2005, 07:06:37 ; Search time 128 Seconds
(without alignments)
191.751 Million cell updates/sec

Title: US-10-018-964-3_COPY_1_15
Perfect score: 87
Sequence: 1 YSGPPGARRNCYE 15

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_P/US10018964/runat_03032005_114707_21385/app_query.fasta_1.199
-DB=issued_Patents_NA -QMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10018964 @CGN 1 105 @runat_03032005_114707_21385 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELACK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	100.0	4086	1 US-08-313-181-1	Sequence 1, Appli
2	63	72.4	1785	5 PCT-US94-12912-1	Sequence 1, Appli
3	63	72.4	3636	1 US-07-753-520B-1	Sequence 1, Appli
4	52	59.8	432	4 US-09-252-991A-2087	Sequence 2087, Ap
5	52	59.8	987	4 US-09-252-991A-1803	Sequence 1803, Ap
6	52	59.8	1272	4 US-09-252-991A-1876	Sequence 1876, Ap
7	51	58.6	303	4 US-09-489-039A-6289	Sequence 6289, Ap
8	51	58.6	1656	4 US-09-902-540-4534	Sequence 4534, Ap
9	51	58.6	10835	4 US-09-902-540-1031	Sequence 1031, Ap
10	50	57.5	202	4 US-09-016-434-285	Sequence 285, App
11	50	57.5	783	4 US-09-902-540-3430	Sequence 3430, Ap
12	50	57.5	18192	4 US-09-902-540-1162	Sequence 1162, Ap

c 13	49	56.3	1348	4	US-09-949-016-4090	Sequence 4090, Ap
c 14	49	56.3	1449	4	US-09-902-540-9457	Sequence 9457, Ap
c 15	49	56.3	13466	4	US-09-902-540-1007	Sequence 1007, Ap
c 16	49	56.3	17723	4	US-09-949-016-15832	Sequence 15832, A
c 17	48	55.2	423	1	US-08-470-179-186	Sequence 186, App
c 18	48	55.2	1209	4	US-09-902-540-7412	Sequence 7412, Ap
c 19	48	55.2	1299	4	US-09-328-352-5	Sequence 5, Appli
c 20	48	55.2	1400	3	US-09-428-589-1	Sequence 1, Appli
c 21	48	55.2	1794	2	US-08-427-497E-5	Sequence 5, Appli
c 22	48	55.2	2600	2	US-08-427-497E-4	Sequence 4, Appli
c 23	48	55.2	3189	2	US-08-427-497E-3	Sequence 3, Appli
c 24	48	55.2	3508	4	US-09-949-016-12325	Sequence 12325, A
c 25	48	55.2	3774	2	US-08-341-843B-1	Sequence 1, Appli
c 26	48	55.2	3774	2	US-08-427-497E-1	Sequence 1, Appli
c 27	48	55.2	3774	2	US-08-427-497E-2	Sequence 2, Appli
c 28	48	55.2	3888	3	US-08-506-296B-13	Sequence 13, Appli
c 29	48	55.2	4781	4	US-09-515-534A-1	Sequence 1, Appli
c 30	48	55.2	6119	4	US-09-902-540-713	Sequence 713, App
c 31	48	55.2	7280	4	US-09-902-540-827	Sequence 827, App
c 32	48	55.2	9562	3	US-08-886-967-1	Sequence 1, Appli
c 33	48	55.2	9562	3	US-09-306-949-1	Sequence 1, Appli
c 34	48	55.2	9562	4	US-09-307-217-1	Sequence 1, Appli
c 35	48	55.2	11575	4	US-09-938-956-5	Sequence 5, Appli
c 36	48	55.2	77536	4	US-09-410-551B-1	Sequence 1, Appli
c 37	48	55.2	77536	4	US-09-940-316B-1	Sequence 1, Appli
c 38	48	55.2	1230025	4	US-09-198-452A-1	Sequence 1, Appli
c 39	48	55.2	1230230	4	US-09-438-185A-1	Sequence 1, Appli
c 40	47	54.0	534	4	US-09-266-965-55	Sequence 55, Appl
c 41	47	54.0	550	4	US-09-621-376-10430	Sequence 10430, A
c 42	47	54.0	601	4	US-09-949-016-85540	Sequence 8540, A
c 43	47	54.0	601	4	US-09-949-016-85541	Sequence 8541, A
c 44	47	54.0	601	4	US-09-949-016-85542	Sequence 8542, A
c 45	47	54.0	954	3	US-09-418-641-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-313-181-1
; Sequence 1, Application US/08313181
; Patent No. 5681735 10201997
; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P.
; APPLICANT: Goldhamer, David J.
; TITLE OF INVENTION: Transcription Control Element for
; Increasing Gene Expression in Myoblasts
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street, Suite 720
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,181
; FILING DATE: 07-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Janet E.
; REGISTRATION NUMBER: 36,252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 563-4100
; TELEFAX: (215) 563-4044
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4086 base pairs
; TYPE: nucleic acid

Not for publication
Copyright

; PRIOR FILING DATE: 2000-07-10
 ; NUMBER OF SEQ ID NOS: 16825
 ; SEQ ID NO 4534
 ; LENGTH: 1656
 ; TYPE: DNA
 ; ORGANISM: Myxococcus xanthus
 US-09-902-540-4534

Alignment Scores:
 Pred. No.: 180 Length: 1656
 Score: 51.00 Matches: 8
 Percent Similarity: 76.92% Conservative: 2
 Best Local Similarity: 61.54% Mismatches: 3
 Query Match: 58.62% Indels: 0
 DB: 4 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x US-09-902-540-4534 (1-1656)

Qy 1 TyrSerGlyProProSerGlyAlaArgArgAsnCys 13
 Db 439 CATCCAGGGCCACCGCTGCCGTAGAGACGGCGTTGT 401

RESULT 9

; Sequence 1031, Application US/09902540
 ; Patent No. 6833447
 ; GENERAL INFORMATION:
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Wiegand, Roger C.
 ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 ; FILE REFERENCE: 38-10(15849)B
 ; CURRENT APPLICATION NUMBER: US/09/902,540
 ; PRIOR FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: 60/217,883
 ; PRIOR FILING DATE: 2000-07-10
 ; NUMBER OF SEQ ID NOS: 16825
 ; SEQ ID NO 1031
 ; LENGTH: 10835
 ; TYPE: DNA
 ; ORGANISM: Myxococcus xanthus
 US-09-902-540-1031

Alignment Scores:
 Pred. No.: 1,44e+03 Length: 10835
 Score: 51.00 Matches: 8
 Percent Similarity: 76.92% Conservative: 2
 Best Local Similarity: 61.54% Mismatches: 3
 Query Match: 58.62% Indels: 0
 DB: 4 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x US-09-902-540-1031 (1-10835)

Qy 1 TyrSerGlyProProSerGlyAlaArgArgAsnCys 13
 Db 3980 CATCCAGGGCCACCGCTGCCGTAGAGACGGCGTTGT 3942

RESULT 10

; Sequence 285, Application US/09016434
 ; Patent No. 6500938
 ; GENERAL INFORMATION:
 ; APPLICANT: Janice Au-Young
 ; APPLICANT: Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
 ; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
 ; NUMBER OF SEQUENCES: 1490
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA

; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/016,434
 ; FILING DATE: HERewith
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0002 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 285:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 202 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: BLADNOT06
 ; CLONE: 1723064
 US-09-016-434-285

Alignment Scores:
 Pred. No.: 24.7 Length: 202
 Score: 50.00 Matches: 8
 Percent Similarity: 81.82% Conservative: 1
 Best Local Similarity: 72.73% Mismatches: 2
 Query Match: 57.47% Indels: 0
 DB: 4 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x US-09-016-434-285 (1-202)

Qy 3 GlyProProSerGlyAlaArgArgAsnCys 13
 Db 132 GGGCCACCTACAGGTTTGAGGAGGAGACCTGC 164

RESULT 11

; Sequence 3430, Application US/09902540
 ; Patent No. 6833447
 ; GENERAL INFORMATION:
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Wiegand, Roger C.
 ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 ; FILE REFERENCE: 38-10(15849)B
 ; CURRENT APPLICATION NUMBER: US/09/902,540
 ; CURRENT FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: 60/217,883
 ; PRIOR FILING DATE: 2000-07-10
 ; NUMBER OF SEQ ID NOS: 16825
 ; SEQ ID NO 3430
 ; LENGTH: 783
 ; TYPE: DNA
 ; ORGANISM: Myxococcus xanthus
 US-09-902-540-3430

Alignment Scores:
 Pred. No.: 111 Length: 783
 Score: 50.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 1

Best Local Similarity: 90.00% Mismatches: 0
Query Match: 57.47% Indels: 0
DB: 4 Gaps: 0
US-10-018-964-3_COPY_1_15 (1-15) x US-09-902-540-3430 (1-783)

Qy 2 SerGlyProSerGlyAlaArgArg 11
Db 91 ACCGGTCCACCTTCTGCTCGCGACGC 62

RESULT 12

US-09-902-540-1162
; Sequence 1162, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1162
; LENGTH: 18192
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1162

Alignment Scores:
Pred. No.: 3.62e+03 Length: 18192
Score: 50.00 Matches: 9
Percent Similarity: 100.00% Conservativeness: 1
Best Local Similarity: 90.00% Mismatches: 0
Query Match: 57.47% Indels: 0
DB: 4 Gaps: 0
US-10-018-964-3_COPY_1_15 (1-15) x US-09-902-540-1162 (1-18192)

Qy 2 SerGlyProSerGlyAlaArgArg 11
Db 16821 ACCGGTCCACCTTCTGCTCGCGACGC 16850

RESULT 13

US-09-949-016-4090/c
; Sequence 4090, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4090
; LENGTH: 1348
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4090

Alignment Scores:
Pred. No.: 286 Length: 1348

Score: 49.00 Matches: 8
Percent Similarity: 81.82% Conservativeness: 1
Best Local Similarity: 72.73% Mismatches: 2
Query Match: 56.32% Indels: 0
DB: 4 Gaps: 0
US-10-018-964-3_COPY_1_15 (1-15) x US-09-949-016-4090 (1-1348)

Qy 3 GlyProSerGlyAlaArgArgAsnCys 13
Db 493 GGACCTCCGTCAGGCTGCTCGAGGAGAGTTCG 461

RESULT 14

US-09-902-540-9457/c
; Sequence 9457, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 9457
; LENGTH: 1449
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-9457

Alignment Scores:
Pred. No.: 310 Length: 1449
Score: 49.00 Matches: 9
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 56.32% Indels: 0
DB: 4 Gaps: 0
US-10-018-964-3_COPY_1_15 (1-15) x US-09-902-540-9457 (1-1449)

Qy 3 GlyProSerGlyAlaArgArg 11
Db 199 GGTCCTCCGTCAGGACCGG 173

RESULT 15

US-09-902-540-1007
; Sequence 1007, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1007
; LENGTH: 13466
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1007

Alignment Scores:
Pred. No.: 3.67e+03 Length: 13466
Score: 49.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 56.32% Indels: 0
 DB: 4 Gaps: 0
 US-10-018-964-3_COPY_1_15 (1-15) x US-09-902-540-1007 (1-13466)
 QY 3 GlyProProSerGlyAlaArgArg 11
 Db 2126 GGTCTCTCGTCGAGCCAGACGG 2152

Search completed: March 7, 2005, 10:06:20
 Job time : 134 secs

BLAST ANALYSIS COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2005, 13:41:45 ; Search time 163 seconds
(without alignments)
35.591 Million cell updates/sec

Title: US-10-018-964-3_COPY_1_15

Perfect score: 87

Sequence: 1 YSGPPSGARRRNCYE 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	87	100.0	320	6	ADA84119 Human MYO
2	87	100.0	320	8	ABM81235 Tumour-as
3	58	66.7	318	3	AAV93769 Amino aci
4	47	54.0	271	3	AB28444 Human fra
5	47	54.0	271	5	ABP62934 Human pro
6	47	54.0	291	3	AAAB56390 Human pol
7	46	52.9	54	4	AAU49323 Propionib
8	46	52.9	54	6	ABM45842 Propionib
9	46	52.9	117	7	ABO61345 Klebsiell
10	45	51.7	1167	6	ABU34694 Protein e
11	45	51.7	1184	6	ABU36648 Protein e
12	45	51.7	1184	7	ADD24621 DNA polym
13	44	50.6	70	4	AAU52239 Propionib
14	44	50.6	70	6	ABM48758 Propionib
15	44	50.6	109	5	ADK34673 Novel hum
16	44	50.6	269	3	AAV74328 Neisseria
17	44	50.6	464	7	ABO77466 Pseudomon
18	43	49.4	94	8	ADQ19641 Human sof
19	43	49.4	140	5	ADK35391 Novel hum
20	43	49.4	184	3	AAV90927 Cenarchae
21	43	49.4	243	6	ABR54306 Human sec
22	43	49.4	258	7	ABO76047 Pseudomon
23	43	49.4	410	3	AAV52590 Human pro
24	43	49.4	475	4	ABG03872 Novel hum
25	42	48.3	42	5	ABG80828 Human SF-

ALIGNMENTS

RESULT 1

ADA84119
ID ADA84119 standard; protein; 320 AA.

XX ADA84119;

XX AC

XX 20-NOV-2003 (first entry)

XX Human MYOD1 protein.

XX human; marker; expressed sequence tag; EST; arabadopsis; tumour;

XX KW stress-induced phenotype; hyperosmotic stress; colon cancer; immunogen;

XX KW vaccine.

XX OS Homo sapiens.

XX PN WO2002103028-A2.

XX PD 27-DEC-2002.

XX PF 30-MAY-2002; 2002WO-IB004189.

XX PR 30-MAY-2001; 2001US-0293999P.

XX PR 22-OCT-2001; 2001US-0330457P.

XX PR 19-FEB-2002; 2002US-0357144P.

XX PA (BIOM-) BIOMEDICAL CENT.

XX PI Baranova AV, Yankovsky NK, Kozlov AP, Lobashev AV, Krukovskaya LL;

XX DR WPI; 2003-175241/17.

XX DR N-PSDB; ADA84118.

XX PT Determining if a nucleic acid is a marker for a phenotype/cell type of interest, by global comparison of expressed sequence tags known to be expressed in the phenotype/cell type with all ESTs expressed in normal tissue.

XX PS Claim 29; Page 499-501; 516pp; English.

XX CC The invention relates to a novel method for determining if a nucleic acid is a marker for a predetermined phenotype/cell type of interest from a biological species. The method comprises performing a global comparison of a group of expressed sequence tags (ESTs) known to be expressed in the phenotype/cell type of interest with all ESTs expressed in normal tissue in order to identify ESTs that are preferentially expressed in the phenotype/cell of interest. A method of the invention is useful for determining whether a nucleic acid is a marker for a predetermined

CC phenotype or cell type of interest from a biological species, preferably
 CC Arabidopsis or human. The cell type of interest is an abnormal cell such
 CC as a tumour cell, and the predetermined phenotype is a stress-induced
 CC phenotype such as hyperosmotic stress or high salt conditions. A method
 CC of the invention is also useful for determining the progression of colon
 CC cancer in a human, for detecting a tumour cell, and for regulating or
 CC preventing the growth of a tumour cell. An antibody of the invention is
 CC useful for detecting the absence or presence of peptides encoded by
 CC tumour-associated markers. A polypeptide of the invention is useful as an
 CC immunogen for vaccinating an animal. The present sequence represents a
 CC tumour-associated antigen of the invention.
 XX
 SQ Sequence 320 AA;

Query Match 100.0%; Score 87; DB 6; Length 320;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSGPPSGARRRNCYE 15
 |||||
 Db 213 YSGPPSGARRRNCYE 227

RESULT 2

ABM81235
 ID ABM81235 standard; protein; 320 AA.

XX
 AC ABM81235;

DT 18-NOV-2004 (first entry)

XX Tumour-associated antigenic target (TAT) polypeptide PRO81964, SEQ:3179.
 XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;
 KW tumour; diagnosis; cell proliferative disorder; breast cancer;
 KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
 KW central nervous system cancer; bladder cancer; pancreatic cancer;
 KW cervical cancer; melanoma; leukaemia; hybridisation probe;
 KW chromosome identification; chromosome mapping; gene mapping;
 KW gene therapy; cytostatic.

XX Homo sapiens.

XX WO2004030615-A2.

XX 15-APR-2004.

XX 29-SEP-2003; 2003WO-US028547.

PR 02-OCT-2002; 2002US-0414971P.

XX (GETH) GENENTECH INC.

XX Wu TD, Zhang Z, Zhou Y;

DR WPI; 2004-347921/32.

DR N-PSDB; ACN39187.

XX New tumor-associated antigenic target polypeptides and nucleic acids,
 PT useful in preparing a medicament for treating or detecting a
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
 PT prostate cancer or tumor.

XX Claim 12; SEQ ID NO 3179; 7273pp; English.

XX The invention relates to human tumour-associated antigenic target (TAT)
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are
 CC overexpressed in cancer tissues compared to normal tissues, and may thus
 CC serve as effective targets for the diagnosis and treatment of cancer in
 CC mammals. The invention also relates to nucleic acid and polypeptide
 CC sequences at least 80% identical to the TAT nucleic acids and
 CC polypeptides; expression vectors and host cells comprising a TAT nucleic
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic

CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
 CC TAT polypeptide; and methods and compositions for the treatment or
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
 CC antibodies, antagonists, binding molecules and compositions are useful
 CC for diagnosing or treating a cell proliferative disorder associated with
 CC increased TAT expression, particularly cancers such as breast cancer,
 CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
 CC cancer, pancreatic cancer, cervical cancer, cancers of the central
 CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
 CC used as hybridisation probes, in chromosome and gene mapping, in
 CC chromosome identification and in gene therapy. The present sequence
 CC represents a TAT polypeptide of the invention
 XX
 SQ Sequence 320 AA;

Query Match 100.0%; Score 87; DB 8; Length 320;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YSGPPSGARRRNCYE 15
 |||||
 Db 213 YSGPPSGARRRNCYE 227

RESULT 3

AY93769
 ID AY93769 standard; protein; 318 AA.

XX
 AC AY93769;

DT 03-OCT-2000 (first entry)

XX Amino acid sequence of the transcription factor protein MyoD.
 DE Transcription factor; MyoD; myoD gene; regulatory gene;
 KW skeletal myogenesis; DNA-binding protein; bHLH domain;
 KW augmentor sequence.

XX Mus sp.

XX Key Location/Qualifiers

FT Domain 1..53 /note= "transcription activation domain"
 FT Domain 10..47 /note= "acidic domain"
 FT Region 100..112 /note= "nuclear localisation signal"
 FT Domain 102..121 /note= "basic domain involved in myogenesis"
 FT Region 122..136 /note= "first helix"
 FT Misc-difference 129 /note= "Pro encoded by TTC"
 FT Region 130..135 /note= "nuclear localisation signal"
 FT Region 137..145 /note= "loop region"
 FT Misc-difference 139 /note= "Asp encoded by AAC"
 FT Misc-difference 141 /note= "Asp encoded by AAC"
 FT Region 146..162 /note= "second helix"
 FT Misc-difference 152 /note= "Asp encoded by AAC"
 FT Misc-difference 223 /note= "Asp encoded by AAT"
 XX WO200037682-A1.
 XX 29-JUN-2000.
 XX 20-DEC-1999; 99WO-US030460.

```
XX 21-DEC-1998; 98US-0113261P.
XX (PROM-) PROMEGA CORP.
XX (SCHE/) SCHENBORN E T.
XX (BRON/) BRONDYK W.
XX (DBER/) DE BERG L G.
XX (ALMO/) ALMOND B D.
XX
XX Schenborn ET, Brondyk W, De Berg LG, Almond BD;
XX
XX WPI; 2000-442694/38.
XX N-PSDB; AAA47075.
XX
XX Improving transient expression of an extrachromosomal foreign gene in
XX cultured eukaryotic cells to produce large amounts of a transgenic gene
XX product in cultured cells.
XX
XX Disclosure; Fig 1B; 133pp; English.
XX
XX The present sequence represents a murine the transcription factor protein
XX MyoD1. It is derived from the myoD1 gene, which is a regulatory gene for
XX skeletal myogenesis. The protein is a DNA-binding protein that binds to
XX the enhancer elements of several muscle-specific genes. The myoD1 gene
XX contains a bHLH domain, which consists of a basic segment that binds to a
XX consensus DNA sequence (CANNTG) element termed the "E-box", and a short
XX alpha helix connected by a loop to a longer alpha helix. The loop is
XX flexible and allows the two helices to fold back against each other. The
XX alpha helix binds to DNA and the HLH structure of another bHLH
XX polypeptide. The second polypeptide can be the same or different, and the
XX resulting protein complexes then bind to DNA in the 5' regulatory regions
XX of genes. An augmentor sequence can be derived from the myoD1 cDNA
XX sequence, which augments post-transfection expression of a foreign gene
XX of interest. The augmentor sequence is used in the method of the
XX invention, for the improving transient gene expression of an
XX extrachromosomal foreign gene in a host cell
XX
XX SQ Sequence 318 AA;
XX
XX Query Match 66.7%; Score 58; DB 3; Length 318;
XX Best Local Similarity 66.7%; Pred. No. 2;
XX Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 YSGPPSGARRRNCYE 15
XX ||||| ||::|
XX 212 YSGPPSGPRRQGDYD 226
XX
XX RESULT 4
XX AAB28444
XX ID AAB28444 standard; protein; 271 AA.
XX
XX AC AAB28444;
XX
XX 31-JAN-2001 (first entry)
XX Human fra-1 protein.
XX
XX Human; fra-1; antisense oligonucleotide; phosphorothioate; cytostatic;
XX antiinflammatory; 2'-methoxyethyl wing; 2'-MOE wing; infection; cancer.
XX
XX Homo sapiens.
XX
XX US6124133-A.
XX
XX 26-SEP-2000.
XX
XX 15-OCT-1999; 99US-00418641.
XX
XX 15-OCT-1999; 99US-00418641.
XX
XX (ISIS-) ISIS PHARM INC.
XX
```

```
PI Taylor JK, Cowsert LM;
XX WPI; 2000-601552/57.
XX N-PSDB; AAC60503.
XX
XX Novel antisense compound 8-30 nucleobases in length targeted to human fra
XX -1 and which specifically hybridizes with and inhibits the expression of
XX human fra-1, useful for modulating the expression of fra-1 in cells.
XX
XX Disclosure; Col 43-46; 38pp; English.
XX
XX The nucleotide sequence encoding this protein was used to design a large
XX number of antisense oligonucleotides which are targeted to nucleic acids
XX encoding fra-1. The sequences may be oligodeoxyribonucleotides or
XX chimeric oligonucleotides containing a central gap region consisting of
XX ten 2'-deoxynucleotides, which is flanked on both sides by 2'-
XX methoxyethyl (2'-MOE) wings. The oligonucleotides have a phosphorothioate
XX backbone and the cytidine residues in the 2'-MOE wings are 5-
XX methylcytidines. The fra-1 antisense oligonucleotides are useful for
XX inhibiting the expression of fra-1 in human cells or tissues. They can be
XX used for diagnostics, therapeutics, prophylaxis and as research reagents
XX and in kits. Use of the antisense compounds may also be useful
XX prophylactically, e.g. to prevent or delay infection, inflammation or
XX tumour formation
XX
XX SQ Sequence 271 AA;
XX
XX Query Match 54.0%; Score 47; DB 3; Length 271;
XX Best Local Similarity 72.7%; Pred. No. 76;
XX Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX QY 3 GPPSGARRRNC 13
XX ||||| |||||
XX Db 87 GPPPGVRRRC 97
XX
XX RESULT 5
XX ABP62934
XX ID ABP62934 standard; protein; 271 AA.
XX
XX AC ABP62934;
XX
XX 14-OCT-2002 (first entry)
XX Human polypeptide SEQ ID NO 371.
XX
XX Human; vulnery; dermatological; neuroprotective; nootropic; cancer;
XX antiparkinsonian; immunostimulant; cytostatic; immunosuppressive;
XX antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;
XX burn; central nervous system disorder; Alzheimer's disease;
XX Parkinson's disease; Huntington's disease; immune disorder;
XX autoimmune disorder; multiple sclerosis; diabetes; allergy.
XX
XX Homo sapiens.
XX
XX WO200218424-A2.
XX
XX 07-MAR-2002.
XX
XX 31-AUG-2001; 2001WO-US027093.
XX
XX 01-SEP-2000; 2000US-00654935.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;
XX Zhao QA, Wang D, Liu C, Drmanac RT, Wehrman T;
XX
XX WPI; 2002-583321/62.
XX N-PSDB; ABQ93413.
XX
XX New polynucleotide and polypeptides, useful for treatment and diagnosis
XX of Alzheimer's, Parkinson's, Huntington's, amyotrophic lateral
XX
```

PT sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple
 PT sclerosis, diabetes and allergies.
 XX
 PS Claim 20; SEQ ID NO 371; 284pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide (I) comprising one of
 CC 245 sequences (AB093288-AB093532). Treating a condition comprising
 CC administering to a mammalian subject a composition comprising the protein
 CC (II) encoded by (I) (ABP62809-ABP63053) or an antibody (III) to (II).
 CC (I), (II) and (III) are useful for diagnostic evaluation of disorders.
 CC (I) is useful for gene therapy of diseases and (II) can be used for
 CC therapeutic treatment. Diseases that may be treated include wound healing
 CC and tissue repair, burns, central nervous system disorders (e.g.
 CC Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral
 CC sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple
 CC sclerosis, diabetes and allergies. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 271 AA;

Query Match 54.0%; Score 47; DB 5; Length 271;
 Best Local Similarity 72.7%; Pred. No. 76;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 3 GPPSGARRRNC 13
 ||| |||||
 Db 87 GPPPGVRRRPC 97

RESULT 6
 AAB56390
 ID AAB56390 standard; protein; 291 AA.
 XX
 AC AAB56390;
 DT 13-MAR-2001 (first entry)
 XX
 DE Human prostate cancer antigen protein sequence SEQ ID NO:968.
 XX
 KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 KW neuroprotective; cytostatic; cardiostatic; immunomodulatory; muscular;
 KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 KW wound; infectious disease.
 XX
 OS Homo sapiens.
 XX
 XX WO200055174-A1.
 XX
 XX 21-SEP-2000.
 XX
 XX 08-MAR-2000; 2000WO-US005988.
 XX
 XX 12-MAR-1999; 99US-0124270P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX (ROSE/) ROSEN C A.
 XX
 XX Rosen CA, Ruben SM;
 PI
 XX WPI; 2000-587513/55.
 DR
 DR N-PSDB; AAF15593.
 XX
 XX Prostate cancer associated gene sequences, referred to as prostate cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as prostate cancer.
 XX
 PS Claim 11; Page 1408-1409; 2338pp; English.
 XX
 CC AAF15566 to AAF16505 encode the human prostate cancer associated

CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
 CC The prostate cancer antigens can have neuroprotective, cytostatic,
 CC cardiostatic, immunomodulatory, muscular, vulnary, gastrointestinal,
 CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
 CC and can be used in gene therapy. The prostate cancer antigen
 CC polynucleotides may be used for detection of prostate cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The prostate cancer antigens may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
 CC AAB57303 represent sequences used in the exemplification of the present
 CC invention
 XX
 SQ Sequence 291 AA;

Query Match 54.0%; Score 47; DB 3; Length 291;
 Best Local Similarity 72.7%; Pred. No. 81;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 3 GPPSGARRRNC 13
 ||| |||||
 Db 107 GPPPGVRRRPC 117

RESULT 7
 AAU49323
 ID AAU49323 standard; protein; 54 AA.
 XX
 AC AAU49323;
 DT 13-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #10219.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 XX WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 XX 20-APR-2001; 2001WO-US012865.
 XX
 XX 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 XX (CORI-) CORIXA CORP.
 XX
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 XX WPI; 2001-616774/71.
 DR
 DR N-PSDB; AAS59545.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX
 XX Example 1; SEQ ID NO 10518; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory

CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of *P. acnes* in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for *P. acnes* proteins. These antibodies can be used to
 CC downregulate expression and activity of *P. acnes* polypeptides and
 CC therefore treat *P. acnes* infections. The antibodies may also be used as
 CC diagnostic agents for determining *P. acnes* presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 54 AA;

Query Match 52.9%; Score 46; DB 4; Length 54;
 Best Local Similarity 66.7%; Pred. No. 23;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GPPSGARRRNCY 14
 |||||
 Db 2 GPPSPRRRGCH 13

RESULT 8
 ABM45842
 ID ABM45842 standard; protein; 54 AA.

AC ABM45842;

XX 20-OCT-2003 (first entry)

XX Propionibacterium acnes predicted ORF-encoded polypeptide #10518.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine.

XX Propionibacterium acnes.

PN WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Vallie-Douglas J;

XX WPI; 2003-381789/36.
 DR N-PSDB; ACF64474.

PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a *P. acnes* protein.

XX Example 1; SEQ ID NO 10518; 1481pp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of *P. acnes* polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a *P. acnes*
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising *P. acnes* polypeptides,

CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of *P. acnes* in a
 CC patient; and a method for inhibiting the development of *P. acnes* in a
 CC patient. The *P. acnes* polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a *P. acnes*
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against *P. acnes*, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the *P. acnes* polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 54 AA;

Query Match 52.9%; Score 46; DB 6; Length 54;
 Best Local Similarity 66.7%; Pred. No. 23;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GPPSGARRRNCY 14
 |||||
 Db 2 GPPSPRRRGCH 13

RESULT 9
 ABO61345

ID ABO61345 standard; protein; 117 AA.

XX ABO61345;

XX 29-JUL-2004 (first entry)

XX Klebsiella pneumoniae polypeptide seqid 7862.

XX Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; vaccine.

XX Klebsiella pneumoniae.

XX US6610836-B1.

XX 26-AUG-2003.

XX 27-JAN-2000; 2000US-00489039.

XX 29-JAN-1999; 99US-0117747P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Breton GL, Osborne M;

XX WPI; 2003-895346/82.
 DR N-PSDB; ACH94896.

XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.

XX Disclosure; SEQ ID NO 7862; 932pp; English.

XX The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention

XX Sequence 117 AA;

Query Match 52.9%; Score 46; DB 7; Length 117;
 Best Local Similarity 66.7%; Pred. No. 48;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 SGPPSGARRRNC 13
 |||:|||||
 Db 1 SGPRAGRRRQC 12

RESULT 10
 ABU34694
 ID ABU34694 standard; protein; 1167 AA.
 AC ABU34694;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #20221.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Mycobacterium bovis.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA38564.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 62618; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC on a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1167 AA;
 Query Match 51.7%; Score 45; DB 6; Length 1167;
 Best Local Similarity 57.1%; Pred. No. 6e+02;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 SGPPSGARRRNCYE 15
 :|||:|||||
 Db 334 AGPPDGYRRAAYE 347

RESULT 11
 ABU36648
 ID ABU36648 standard; protein; 1184 AA.
 XX
 AC ABU36648;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #22175.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA40518.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 64572; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC on a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required

CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1184 AA;

Query Match 51.7%; Score 45; DB 6; Length 1184;
Best Local Similarity 57.1%; Pred. No. 6.1e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Oy 2 SGPPSGARRNCYE 15
:|||||
Db 335 AGPPDGYRRAAYE 348

RESULT 12

ADD24621
ID ADD24621 standard; protein; 1184 AA.

XX AC ADD24621;

XX DT 15-JAN-2004 (first entry)

XX DE DNA polymerase III subunit related protein, SEQ ID NO 140.

XX KW DNA polymerase III replicase; dnaE; hola; holB; holC; hold; hola; dnaX;
XX dnaN; SSB; dnaG; dnaQ; bacterial DNA replication; bacterial infection.

XX OS Mycobacterium tuberculosis.

XX PN WO200292769-A2.

XX PD 21-NOV-2002.

XX PF 14-MAY-2002; 2002WO-US015111.

XX PR 14-MAY-2001; 2001US-0290725P.

XX PR 05-NOV-2001; 2001US-0332644P.

XX PA (REPL-) REPLIDYNE INC.

XX PI Bullard JM, Janjic N, Mchenry CS;

XX WPI; 2003-120672/11.

XX Screening for a compound that modulates the activity of *Y. pestis* and/or
XX *P. aeruginosa* DNA polymerase III replicase, useful for reducing bacterial
XX DNA replication and infection in animals, plants, humans and surrounding
XX environment.

XX Disclosure; SEQ ID NO 140; 389pp; English.

XX The invention relates to a novel screening method for a compound that
XX modulates the activity of a DNA polymerase III replicase. The novel
XX method comprises contacting an isolated replicase with at least one test
XX compound under conditions permissive for replicase activity, and
XX comparing the activity of the replicase in the presence and absence of
XX the test compound, where a change in the activity of the replicase

CC indicates a compound that modulates the activity of the replicase. The
CC replicase comprises an isolated nucleic acid from *Yersinia pestis* or a
CC *Pseudomonas aeruginosa* encoding a DNA polymerase III subunit protein. The
CC DNA polymerase III subunit proteins include dnaE, hola, holB, holC, hold,
CC holaE, dnaX, dnaN, SSB, dnaG, and dnaQ. The methods and compositions of
CC the present invention are useful for reconstituting replicases and
CC polymerases for sequencing, amplification and screening for compounds
CC which modulate the function of the polymerase or replicase particularly
CC in bacterial DNA replication, thereby reducing bacterial infection in
CC animals, plants, humans and the surrounding environment. This sequence
CC represents a DNA polymerase III subunit related protein of the invention.

XX SQ Sequence 1184 AA;

Query Match 51.7%; Score 45; DB 7; Length 1184;
Best Local Similarity 57.1%; Pred. No. 6.1e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Oy 2 SGPPSGARRNCYE 15
:|||||
Db 335 AGPPDGYRRAAYE 348

RESULT 13

AAU52239
ID AAU52239 standard; protein; 70 AA.

XX AC AAU52239;

XX DT 27-FEB-2002 (first entry)

XX DE Propionibacterium acnes immunogenic protein #13135.

XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.

XX OS Propionibacterium acnes.

XX PN WO200181581-A2.

XX PD 01-NOV-2001.

XX PF 20-APR-2001; 2001WO-US012865.

XX PR 21-APR-2000; 2000US-0199047P.

XX PR 02-JUN-2000; 2000US-0208841P.

XX PR 07-JUL-2000; 2000US-0216747P.

XX PA (CORI-) CORIXA CORP.

XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AAS59553.

XX Propionibacterium acnes polypeptides and nucleic acids useful for

XX vaccinating against and diagnosing infections, especially useful for

XX treating acne vulgaris.

XX Example 1; SEQ ID NO 13434; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX *P. acnes*. The disorders include SAPHO syndrome (synovitis, acne,
XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX *P. acnes* is also involved in infections of bone, joints and the central
XX nervous system, however it is particularly involved in the inflammatory
XX lesions associated with acne vulgaris. A method for detecting the
XX presence or absence of *P. acnes* in a patient comprises contacting a

CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 70 AA;
 Query Match 50.6%; Score 44; DB 4; Length 70;
 Best Local Similarity 72.7%; Pred. No. 59;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SGPPSGARRRN 12
 Db 54 SGPPKARRRS 64
 RESULT 14
 ABM48758
 ID ABM48758 standard; protein; 70 AA.
 AC ABM48758;
 XX
 XX 20-OCT-2003 (first entry)
 DT
 XX
 DE Propionibacterium acnes predicted ORF-encoded polypeptide #13434.
 XX
 XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine.
 KW
 XX
 OS Propionibacterium acnes.
 XX
 XX WO2003033515-A1.
 PN
 XX
 XX 24-APR-2003.
 PD
 XX 11-OCT-2002; 2002WO-US032727.
 PF
 XX 15-OCT-2001; 2001US-00978825.
 PR
 XX (CORI-) CORIXA CORP.
 PA
 XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Valliave-Douglas J;
 XX
 XX WPI; 2003-381789/36.
 DR N-PSDB; ACF64482.
 DR
 XX
 XX

PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 PS Example 1; SEQ ID NO 13434; 1481pp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit

CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 70 AA;
 Query Match 50.6%; Score 44; DB 6; Length 70;
 Best Local Similarity 72.7%; Pred. No. 59;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SGPPSGARRRN 12
 Db 54 SGPPKARRRS 64
 RESULT 15
 ADK34673
 ID ADK34673 standard; protein; 109 AA.
 XX
 AC ADK34673;
 XX
 XX 06-MAY-2004 (first entry)
 DT
 XX
 DE Novel human polypeptide SeqID6755.
 XX
 XX antiarthritic; antiparkinsonian; neuroprotective; nootropic;
 KW immunosuppressive; cytostatic; antipsoriatic; antiinflammatory;
 KW antibacterial; antiviral; antifungal; antiparasitic; gene therapy;
 KW arthritis; Parkinson's; Alzheimer's; autoimmune disease; cancer;
 KW psoriasis; inflammatory bowel disease; infection; bacteria; virus;
 KW fungus; parasite; human.
 XX
 OS Homo sapiens.
 XX

XX
 FH Key Location/Qualifiers
 FT Misc-difference 1..109
 FT /label= OTHER
 FT /note= "OTHER= All Xaa's in this sequence are unknown
 FT amino acids or the site of a stop codon within the DNA
 FT sequence"
 XX
 XX WO200216439-A2.
 FN
 XX
 XX 28-FEB-2002.
 PD
 XX
 XX 05-MAR-2001; 2001WO-US004941.
 PF
 XX
 XX 07-MAR-2000; 2000US-00519705.
 PR 19-MAY-2000; 2000US-00574454.
 PR (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Drmanac RT;
 XX
 XX WPI; 2002-280918/32.
 DR

XX Isolated polynucleotide encoding bone marrow derived polypeptides useful
 PT for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis, Crohn's
 PT disease, and inflammatory bowel disease.
 XX
 XX Claim 20; SEQ ID NO 6755; 504pp; English.

XX This invention relates to a novel isolated polynucleotide comprising a
CC nucleotide sequence selected from one of 1680 sequences, a mature protein
CC coding portion of them, an active domain of them and their complementary
CC sequences. The invention may be useful for the production of compounds
CC with an antiarthritic, antiparkinsonian, neuroprotective, nootropic,
CC immunosuppressive, cytostatic, antipsoriatic, antiinflammatory,
CC antibacterial, antiviral, antifungal or antiparasitic activity. In
CC addition, the disclosed sequences may be useful for gene therapy. The
CC polypeptides or their antibodies are useful for treating many diseases
CC such as arthritis, Parkinson's, Alzheimer's, autoimmune diseases, cancer,
CC psoriasis, inflammatory bowel disease and infections caused by bacteria,
CC viruses, fungi or parasites. The present sequence is that of a human
XX polypeptide of the invention.

SQ Sequence 109 AA;

Query Match 50.6%; Score 44; DB 5; Length 109;
Best Local Similarity 63.6%; Pred. No. 89;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GPPSGARRRNC 13
|||.|||
Db 86 GPPFGGGRGNC 96

Search completed: March 3, 2005, 13:56:27
Job time : 164 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2005, 13:51:16 ; Search time 42 Seconds
(without alignments)
26.660 Million cell updates/sec

Title: US-10-018-964-3_COPY_1_15

Perfect score: 87

Sequence: 1 YSGPPSGARRNCYE 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	52.9	117	4	US-09-489-039A-7862
2	44	50.6	464	4	US-09-252-991A-26212
3	43	49.4	184	4	US-09-408-020-32
4	43	49.4	258	4	US-09-252-991A-24793
5	43	49.4	346	1	US-07-951-715A-19
6	43	49.4	346	2	US-08-459-448A-19
7	43	49.4	346	3	US-08-459-595A-19
8	43	49.4	346	3	US-08-459-504B-19
9	43	49.4	346	3	US-08-459-444-19
10	43	49.4	346	3	US-09-547-422-19
11	43	49.4	346	4	US-09-988-452-19
12	43	49.4	410	3	US-09-083-521-2
13	42	48.3	276	4	US-09-252-991A-18128
14	42	48.3	314	3	US-08-927-219-6
15	42	48.3	613	4	US-09-252-991A-22424
16	42	48.3	672	4	US-09-252-991A-18930
17	41	47.1	127	4	US-09-252-991A-26884
18	41	47.1	145	4	US-09-252-991A-21532
19	41	47.1	146	4	US-09-252-991A-18183
20	41	47.1	156	4	US-09-252-991A-21289
21	41	47.1	260	4	US-09-252-991A-25992
22	41	47.1	341	4	US-09-252-991A-27955
23	41	47.1	373	4	US-09-252-991A-29008
24	41	47.1	690	4	US-09-252-991A-16715
25	40.5	46.6	598	4	US-09-854-845-41
26	40.5	46.6	603	4	US-09-854-845-39
27	40.5	46.6	697	4	US-09-854-845-25

```

28 40.5 46.6 702 4 US-09-854-845-23 Sequence 23, Appl
29 40.5 46.6 739 4 US-09-854-845-45 Sequence 45, Appl
30 40.5 46.6 744 4 US-09-854-845-43 Sequence 43, Appl
31 40.5 46.6 766 4 US-09-854-845-49 Sequence 49, Appl
32 40.5 46.6 771 4 US-09-854-845-47 Sequence 47, Appl
33 40.5 46.6 838 4 US-09-854-845-29 Sequence 29, Appl
34 40.5 46.6 843 4 US-09-854-845-27 Sequence 27, Appl
35 40.5 46.6 865 4 US-09-854-845-33 Sequence 33, Appl
36 40.5 46.6 870 4 US-09-854-845-31 Sequence 31, Appl
37 40 46.0 158 4 US-09-252-991A-31073 Sequence 31073, A
38 40 46.0 186 4 US-09-248-796A-15677 Sequence 15677, A
39 40 46.0 225 4 US-09-252-991A-29348 Sequence 29348, A
40 40 46.0 273 4 US-09-252-991A-31733 Sequence 31733, A
41 40 46.0 331 4 US-09-252-991A-19406 Sequence 19406, A
42 40 46.0 369 4 US-09-252-991A-25394 Sequence 25394, A
43 40 46.0 430 4 US-09-252-991A-25820 Sequence 25820, A
44 40 46.0 526 4 US-09-925-637-74 Sequence 74, Appl
45 40 46.0 622 4 US-09-252-991A-32308 Sequence 32308, A

```

ALIGNMENTS

```

RESULT 1
US-09-489-039A-7862
; Sequence 7862, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7862
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7862

```

```

Query Match 52.9%; Score 46; DB 4; Length 117;
Best Local Similarity 66.7%; Pred. No. 6.8;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 2 SGPPSGARRNC 13
   |||:|||||
Db 1 SGPRAGGRRQC 12

```

```

RESULT 2
US-09-252-991A-26212
; Sequence 26212, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26212
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26212

```

```

Query Match      50.6%; Score 44; DB 4; Length 464;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 GPPSGARR 10
      |||||
Db      101 GPPSGARR 108

RESULT 3
US-09-408-020-32
; Sequence 32, Application US/09408020
; Patent No. 6632937
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCOEP 002A
; CURRENT APPLICATION NUMBER: US/09/408,020
; CURRENT FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/102,294
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-09-408-020-32

Query Match      49.4%; Score 43; DB 4; Length 184;
Best Local Similarity 53.3%; Pred. No. 31;
Matches 8; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

QY      1 YSGPPSGARRRNCYE 15
      |.|.|.|.|.|.|.|
Db      54 YNGTTPGV--KNCYE 66

RESULT 4
US-09-252-991A-24793
; Sequence 24793, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24793
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24793

Query Match      49.4%; Score 43; DB 4; Length 258;
Best Local Similarity 72.7%; Pred. No. 43;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      3 GPPSGARRRNC 13
      |||||
Db      139 GLPSSGRRRRC 149

RESULT 5

```

```

US-07-951-715A-19
; Sequence 19, Application US/07951715A
; Patent No. 5625136
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Lannis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/951,715A
; FILING DATE: 25-SEP-1992
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: S-18805/A/CGC 1577/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8615
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 346 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-951-715A-19

Query Match      49.4%; Score 43; DB 1; Length 346;
Best Local Similarity 42.9%; Pred. No. 57;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      2 SGPPSGARRRNCYE 15
      |.|.|.|.|.|.|
Db      63 AAPQAGRRRRCHQ 76

RESULT 6
US-08-459-448A-19
; Sequence 19, Application US/08459448A
; Patent No. 5859336
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.

```


APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.
APPLICANT: Evola, Stephen V.
APPLICANT: Crossland, Lyle D.
APPLICANT: Wright, Martha S.
APPLICANT: Merlin, Ellis J.
APPLICANT: Launis, Karen L.
APPLICANT: Rothstein, Steven J.
APPLICANT: Bowman, John L.
APPLICANT: Dawson, Cindy G.
APPLICANT: Dunder, Erik M.
APPLICANT: Pace, Gary M.
APPLICANT: Suttie, Janet L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 585936artis Corporation
STREET: Patent & Trademark Dept., 520 White Plains
STREET: Rd., POB 2005
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459.448A
FILING DATE: 02-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: 40403
REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV4
TELEPHONE: (919)541-8582
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 346 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-459-448A-19
Query Match 49.4%; Score 43; DB 2; Length 346;
Best Local Similarity 42.9%; Pred. No. 57;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Qy 2 SGPPSGARRRNCYE 15
Db 63 AAPPGARRRRCHQ 76
RESULT 7
US-08-459-595A-19
Sequence 19, Application US/08459595A
Patent No. 6018104
GENERAL INFORMATION:
APPLICANT: Koziel, Michael G.
APPLICANT: Desai, Nalini M.
APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.

APPLICANT: Warren, Gregory W.
APPLICANT: Evola, Stephen V.
APPLICANT: Crossland, Lyle D.
APPLICANT: Wright, Martha S.
APPLICANT: Merlin, Ellis J.
APPLICANT: Launis, Karen L.
APPLICANT: Rothstein, Steven J.
APPLICANT: Bowman, John L.
APPLICANT: Dawson, Cindy G.
APPLICANT: Dunder, Erik M.
APPLICANT: Pace, Gary M.
APPLICANT: Suttie, Janet L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6018104artis Corporation
STREET: Patent & Trademark Dept., 520 White Plains
STREET: Rd., POB 2005
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459.595A
FILING DATE: 02-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: 40403
REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV3
TELEPHONE: (919)541-8582
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 346 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-459-595A-19
Query Match 49.4%; Score 43; DB 3; Length 346;
Best Local Similarity 42.9%; Pred. No. 57;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Qy 2 SGPPSGARRRNCYE 15
Db 63 AAPPGARRRRCHQ 76
RESULT 8
US-08-459-504B-19
Sequence 19, Application US/08459504B
Patent No. 6075185
GENERAL INFORMATION:
APPLICANT: Koziel, Michael G.
APPLICANT: Desai, Nalini M.
APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.

APPLICANT: Crossland, Lyle D.
APPLICANT: Wright, Martha S.
APPLICANT: Merlin, Ellis J.
APPLICANT: Launis, Karen L.
APPLICANT: Rothstein, Steven J.
APPLICANT: Bowman, Cindy G.
APPLICANT: Dawson, John L.
APPLICANT: Dunder, Erik M.
APPLICANT: Pace, Gary M.
APPLICANT: Suttie, Janet L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6075185artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,504B
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/459,595
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC1577/CIP/DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 346 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-459-504B-19

Query Match 49.4%; Score 43; DB 3; Length 346;
Best Local Similarity 42.9%; Pred. No. 57;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 SGPPSGARRRNCYE 15
Db 63 AAPQAGRRRRCHQ 76

RESULT 9
US-08-459-444-19
Sequence 19, Application US/08459444A
Patent No. 6121014
GENERAL INFORMATION:
APPLICANT: Koziel, Michael G.
Desai, Nalini M.
Lewis, Kelly S.
Kramer, Vance C.
Warren, Gregory W.
Evola, Stephen V.
Crossland, Lyle D.

Wright, Martha S.
Merlin, Ellis J.
Launis, Karen L.
TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED
NUCLEIC ACID CODING SEQUENCE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6121014artis Agribusiness Biotechnology Research, Inc.
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,444A
FILING DATE: 02-JUN-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-18805/P1/CGC1577/CIP/DIV6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 346 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-08-459-444-19

Query Match 49.4%; Score 43; DB 3; Length 346;
Best Local Similarity 42.9%; Pred. No. 57;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 SGPPSGARRRNCYE 15
Db 63 AAPQAGRRRRCHQ 76

RESULT 10
US-09-547-422-19
Sequence 19, Application US/09547422
Patent No. 6320100
GENERAL INFORMATION:
APPLICANT: Koziel, Michael G.
Desai, Nalini M.
Lewis, Kelly S.
Kramer, Vance C.
Warren, Gregory W.
Evola, Stephen V.
Crossland, Lyle D.
Wright, Martha S.
Merlin, Ellis J.
Launis, Karen L.

TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6320100artis Agribusiness Biotechnology Research, Inc.
STREET: 3054 Cornwallis Road

CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/547,422
FILING DATE: 11-APR-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/459,595
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-18805H
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 346 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-547-422-19
Query Match 49.4%; Score 43; DB 3; Length 346;
Best Local Similarity 42.9%; Pred. No. 57;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Qy 2 SGPPSGARRRNCYE 15
Db 63 AAPPOAGRRRRCHQ 76
RESULT 11
US-09-988-462-19
Sequence 19, Application US/09988462
Patent No. 6720488
GENERAL INFORMATION:
APPLICANT: Koziel, Michael G.
Desai, Nalini M.
Lewis, Kelly S.
Kramer, Vance C.
Warren, Gregory W.
Evola, Stephen V.
Crossland, Lyle D.
Wright, Martha S.
Merlin, Ellis J.
Launis, Karen L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: Syngenta Biotechnology, Inc.
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/988,462
FILING DATE: 20-NOV-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/547,422
FILING DATE: 11-APR-2000
APPLICATION NUMBER: US 08/459,504
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-18805I
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 346 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-988-462-19
Query Match 49.4%; Score 43; DB 4; Length 346;
Best Local Similarity 42.9%; Pred. No. 57;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Qy 2 SGPPSGARRRNCYE 15
Db 63 AAPPOAGRRRRCHQ 76
RESULT 12
US-09-083-521-2
Sequence 2, Application US/09083521
Patent No. 6048970
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,521
FILING DATE: Herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0527 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166

```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTUT03
; CLONE: 1999442
US-09-083-521-2

Query Match 49.4%; Score 43; DB 3; Length 410;
Best Local Similarity 72.7%; Pred. No. 68;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GPPSGARRNC 13
Db 190 GVPSGARARGC 200

RESULT 13
US-09-252-991A-18128
; Sequence 18128, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18128
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18128

Query Match 48.3%; Score 42; DB 4; Length 276;
Best Local Similarity 66.7%; Pred. No. 65;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SGPGTARRRNC 13
Db 182 SGPGTARRRNC 193

RESULT 14
US-08-927-219-6
; Sequence 6, Application US/08927219
; Patent No. 6187533
; GENERAL INFORMATION:
; APPLICANT: Bell, Graeme I.
; APPLICANT: Yamagata, Kazuya
; APPLICANT: Oda, Naohisa
; APPLICANT: Kaisaki, Pamela J.
; APPLICANT: Furuta, Hiroto
; APPLICANT: Horikawa, Yukio
; APPLICANT: Menzel, Stephen
; TITLE OF INVENTION: MUTATIONS IN THE DIABETES SUSCEPTIBILITY
; TITLE OF INVENTION: GENES HEPATOCYTE NUCLEAR FACTOR (HNF) 1 ALPHA, HNF-1BETA
; TITLE OF INVENTION: AND HNF-4ALPHA
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA

```

```

; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,219
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,679
; FILING DATE: 30-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/028,056
; FILING DATE: 02-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/025,719
; FILING DATE: 10-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: ARCD:272
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 314 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-927-219-6

Query Match 48.3%; Score 42; DB 3; Length 314;
Best Local Similarity 61.5%; Pred. No. 74;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 YSGPPSGARRNC 13
Db 286 YSGPPPRARPGTC 298

RESULT 15
US-09-252-991A-22424
; Sequence 22424, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22424
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22424

Query Match 48.3%; Score 42; DB 4; Length 613;
Best Local Similarity 58.3%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 SGPPSGARRNC 13
Db 597 SATPGPRRNC 608

```

Search completed: March 3, 2005, 14:00:55
Job time : 42 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2005, 13:57:17 ; Search time 134 Seconds
(without alignments)
36.833 Million cell updates/sec

Title: US-10-018-964-3_COPY_1_15

Perfect score: 87
Sequence: 1 YSGPPSGARRNCYE 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1391452 seqs, 329044822 residues

Total number of hits satisfying chosen parameters: 1391452

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pcp.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pcp.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pcp.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	100.0	320	14	US-10-157-031-401
2	49	56.3	278	15	US-10-424-599-245359
3	48	55.2	1278	16	US-10-437-963-126355
4	47	54.0	271	15	US-10-363-616-371
5	47	54.0	291	9	US-09-925-300-968
6	47	54.0	316	16	US-10-437-963-172815
7	46	52.9	259	15	US-10-424-599-170072
8	45	51.7	336	15	US-10-425-114-60305
9	45	51.7	1167	15	US-10-282-122A-62618
10	45	51.7	1184	15	US-10-282-122A-64572
11	44	50.6	197	15	US-10-425-114-48971
12	44	50.6	103	15	US-10-424-599-203545
13	44	50.6	289	15	US-10-425-114-40066

14	43.5	50.0	73	15	US-10-424-599-235980	Sequence 235980,
15	43	49.4	46	15	US-10-424-599-215744	Sequence 215744,
16	43	49.4	128	15	US-10-424-599-177399	Sequence 177399,
17	43	49.4	155	16	US-10-437-963-109816	Sequence 109816,
18	43	49.4	184	13	US-10-027-806-32	Sequence 32, Appl
19	43	49.4	184	13	US-10-034-623-32	Sequence 32, Appl
20	43	49.4	184	14	US-10-027-801-32	Sequence 32, Appl
21	43	49.4	184	14	US-10-029-120-32	Sequence 19, Appl
22	43	49.4	346	10	US-09-988-462-19	Sequence 2, Appli
23	43	49.4	410	9	US-09-963-896-2	Sequence 36762, A
24	43	49.4	433	15	US-10-425-114-36762	Sequence 46439, A
25	42	48.3	66	9	US-09-864-761-46439	Sequence 189306,
26	42	48.3	98	16	US-10-437-963-189306	Sequence 4, Appli
27	42	48.3	101	9	US-09-816-279-4	Sequence 172573,
28	42	48.3	106	16	US-10-437-963-172573	Sequence 153669,
29	42	48.3	148	16	US-10-437-963-203562	Sequence 203562,
30	42	48.3	157	16	US-10-437-963-106027	Sequence 106027,
31	42	48.3	163	16	US-10-437-963-106027	Sequence 46434, A
32	42	48.3	227	15	US-10-425-114-46434	Sequence 53441, A
33	42	48.3	230	15	US-10-425-114-53441	Sequence 55977, A
34	42	48.3	296	15	US-10-425-114-55977	Sequence 163320,
35	42	48.3	297	16	US-10-437-963-163320	Sequence 6, Appli
36	42	48.3	314	10	US-09-754-106-6	Sequence 64844, A
37	42	48.3	498	15	US-10-425-114-64844	Sequence 73018, A
38	42	48.3	594	15	US-10-425-114-73018	Sequence 104208,
39	41.5	47.7	319	16	US-10-437-963-104208	Sequence 270657,
40	41	47.1	63	15	US-10-424-599-270657	Sequence 30676, A
41	41	47.1	66	14	US-10-029-386-30676	Sequence 277397,
42	41	47.1	86	15	US-10-424-599-277397	Sequence 58129, A
43	41	47.1	113	16	US-10-767-701-58129	Sequence 148747,
44	41	47.1	132	16	US-10-437-963-148747	Sequence 232074,
45	41	47.1	141	15	US-10-424-599-232074	

ALIGNMENTS

RESULT 1

US-10-157-031-401
; Sequence 401, Application US/10157031
; Publication No. US2003010890A1
; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Yankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobashev, A. V.
; APPLICANT: Krukovskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
; FILE REFERENCE: 2760-103
; CURRENT APPLICATION NUMBER: US/10/157,031
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 401
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-157-031-401

Query Match 100.0%; Score 87; DB 14; Length 320;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YSGPPSGARRNCYE 15

Db 213 YSGPPSGARRNCYE 227

RESULT 2

US-10-424-599-245359
; Sequence 245359, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 245359
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(278)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_63590C.1.pep
US-10-424-599-245359

Query Match 56.3%; Score 49; DB 15; Length 278;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 PPSGARRRNCYE 15
||| | | | |
Db 170 PPSGVKIECYE 181

RESULT 3
US-10-437-963-126355
; Sequence 126355, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 126355
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1278)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_2890C.1.pep
US-10-437-963-126355

Query Match 55.2%; Score 48; DB 16; Length 1278;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PPSGARRRNC 13
||| | | | |
Db 12 PPGARRRRC 21

RESULT 4
US-10-363-616-371

; Sequence 371, Application US/10363616
; Publication No. US20040044181A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-113 (793)
; CURRENT APPLICATION NUMBER: US/10/363,616
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 09/654,935
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 490
; SEQ ID NO 371
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-363-616-371

Query Match 54.0%; Score 47; DB 15; Length 271;
Best Local Similarity 72.7%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GPPSGARRRNC 13
||| | | | |
Db 87 GPPGVRRRPC 97

RESULT 5
US-09-925-300-968
; Sequence 968, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 968
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-968

Query Match 54.0%; Score 47; DB 9; Length 291;
Best Local Similarity 72.7%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GPPSGARRRNC 13
||| | | | |
Db 107 GPPGVRRRPC 117

RESULT 6
US-10-437-963-172815
; Sequence 172815, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 126355
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1278)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_2890C.1.pep
US-10-437-963-172815

RESULT 4
US-10-363-616-371

; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 172815
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(316)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_70914C.1.pap
US-10-437-963-172815

Query Match 54.0%; Score 47; DB 16; Length 316;
Best Local Similarity 61.5%; Pred. No. 38;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 YSGPPSGARRRNC 13
: ||| ||| |
Db 46 FSPPPGPRRRRC 58

RESULT 7
US-10-424-599-170072
; Sequence 170072, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 170072
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_124590C.1.pap
US-10-424-599-170072

Query Match 52.9%; Score 46; DB 15; Length 259;
Best Local Similarity 63.6%; Pred. No. 45;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PPSGARRRNCY 14
||| : |||
Db 78 PPSGRRNCY 88

RESULT 8
US-10-425-114-60305
; Sequence 60305, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53213)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60305
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3606-028-E11_FLI.pap
US-10-425-114-60305

Query Match 51.7%; Score 45; DB 15; Length 336;
Best Local Similarity 70.0%; Pred. No. 82;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PSGARRRNCY 14
: ||| |||
Db 88 PADARRRNCY 97

RESULT 9
US-10-282-122A-62618
; Sequence 62618, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62618
; LENGTH: 1167
; TYPE: PRT
; ORGANISM: Mycobacterium bovis
US-10-282-122A-62618

Query Match 51.7%; Score 45; DB 15; Length 1167;
Best Local Similarity 57.1%; Pred. No. 2.7e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 SGPPSGARRRNCYE 15

```
Db      334 AGPPDGYRERAAAYE 347
      :|||:|:|:|:|:|
RESULT 10
US-10-282-122A-64572
; Sequence 64572, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2003-02-20
; PRIOR FILING DATE: 60/191,078
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64572
; LENGTH: 1184
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64572

Query Match      51.7%; Score 45; DB 15; Length 1184;
Best Local Similarity 57.1%; Pred. No. 2.7e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      2 SGPPSGARRRNCYE 15
Db      335 AGPPDGYRERAAAYE 348
      :|||:|:|:|:|:|
RESULT 11
US-10-425-114-48971
; Sequence 48971, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 48971
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700429912_FLI.pep
US-10-425-114-48971

Query Match      50.6%; Score 44; DB 15; Length 97;
Best Local Similarity 63.6%; Pred. No. 35;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      3 GPPSGARRRNC 13
Db      1 GPPPASRRRTC 11
      :|||:|:|:|:|
RESULT 12
US-10-424-599-203545
; Sequence 203545, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 203545
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_25827C.1.pep
US-10-424-599-203545

Query Match      50.6%; Score 44; DB 15; Length 103;
Best Local Similarity 58.3%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      4 PPSGARRRNCYE 15
Db      27 PPSGVHKIECYE 38
      :|||:|:|:|:|
RESULT 13
US-10-425-114-40066
; Sequence 40066, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
```

```
; SEQ ID NO 40066
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700981026_FLI.pep
US-10-425-114-40066

Query Match          50.6%; Score 44; DB 15; Length 289;
Best Local Similarity 58.3%; Pred. No. 1e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy      4 PPSGARRRNCYE 15
Db      193 PPSGVHKIECYE 204
      ||||| : |||
      ||||| : |||

RESULT 14
US-10-424-599-235980
; Sequence 235980, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 235980
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_55117C.1.pep
US-10-424-599-235980

Query Match          50.0%; Score 43.5; DB 15; Length 73;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy      1 YSGPPSG-ARRRNCY 14
Db      39 YSGGRSSRCRRNCY 53
      ||||| : |||||
      ||||| : |||||

RESULT 15
US-10-424-599-215744
; Sequence 215744, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 215744
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_36846C.1.pep
US-10-424-599-215744

Query Match          49.4%; Score 43; DB 15; Length 46;
```

```
Best Local Similarity 54.5%; Pred. No. 25;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy      4 PPSGARRRNCY 14
Db      34 PTSGSRKKSCY 44
      ||||| : |||||
      ||||| : |||||
```

Search completed: March 3, 2005, 14:12:01
Job time : 135 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 7, 2005, 09:09:13 ; Search time 492 Seconds
(without alignments)
180.910 Million cell updates/sec

Title: US-10-018-964-3_COPY_1_15

Perfect score: 87
Sequence: 1 YSGPPSGARRNCYE 15

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5401638 seqs, 2966923429 residues

Total number of hits satisfying chosen parameters: 10803276

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPRO_spool_p/US10018964/runat_03032005_114709_21446/app_query.fasta_1.199
-DB=Published Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pcr -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -WMAP=US10018964 @cgn1_1_723 @runat_03032005_114709_21446
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT_DSBLCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	100.0	2480	15	US-10-157-031-400
2	87	100.0	2744	18	US-10-723-860-7932
3	87	100.0	3025	18	US-10-473-126-45
4	64	73.6	3025	18	US-10-473-126-183
5	56	64.4	3025	18	US-10-473-126-184
6	53	60.9	461	18	US-10-425-115-124483
7	52	59.8	542	18	US-10-357-930-12348
8	52	59.8	2358	17	US-10-282-122A-13812
9	51	58.6	855	15	US-10-128-714-2252
10	51	58.6	855	15	US-10-128-714-7252
11	51	58.6	908	15	US-10-128-714-1252
12	51	58.6	908	15	US-10-128-714-6252
13	51	58.6	2328	18	US-10-437-963-48056
14	51	58.6	2908	15	US-10-128-714-252
15	51	58.6	2908	15	US-10-128-714-5252
16	50	57.5	202	17	US-10-305-720-285
17	50	57.5	356	18	US-10-425-115-53450
18	50	57.5	400	18	US-10-425-115-113349
19	50	57.5	754	18	US-10-437-963-65811
20	50	57.5	768	18	US-10-425-115-75153
21	50	57.5	5023	18	US-10-723-860-376
22	50	57.5	5084	18	US-10-335-053-96
23	50	57.5	5117	18	US-10-723-860-5073
24	50	57.5	22017	19	US-10-211-028-12
25	50	57.5	52302	11	US-09-997-722-4
26	50	57.5	90597	11	US-10-211-028-1
27	49	56.3	426	17	US-10-424-599-129806
28	49	56.3	520	18	US-10-767-701-15225
29	49	56.3	687	17	US-10-425-115-102582
30	49	56.3	799	17	US-10-425-114-3077
31	49	56.3	1012	18	US-10-437-963-46584
32	49	56.3	1048	18	US-10-437-963-46585
33	49	56.3	1055	17	US-10-236-392-9
34	49	56.3	1300	9	US-09-880-107-3731
35	49	56.3	1309	9	US-09-880-107-2235
36	49	56.3	1309	17	US-10-236-392-5
37	49	56.3	1309	18	US-10-283-975A-169
38	49	56.3	1430	17	US-10-062-674-1931
39	49	56.3	1745	17	US-10-260-238-4156
40	49	56.3	1783	17	US-10-424-599-102517
41	49	56.3	1887	15	US-10-156-761-7126
42	49	56.3	1908	18	US-10-437-963-55618
43	49	56.3	2124	18	US-10-425-115-41288
44	49	56.3	2503	18	US-10-437-963-35317
45	49	56.3	9762	19	US-10-762-107-73

ALIGNMENTS

RESULT 1

US-10-157-031-400
; Sequence 400, Application US/10157031
; Publication No. US20030108890A1
; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Yankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobashev, A. V.
; APPLICANT: Krutovskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
; FILE REFERENCE: 2760-103
; CURRENT APPLICATION NUMBER: US/10/157.031
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 400
; LENGTH: 2480
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-157-031-400

Alignment Scores:
Pred. No.: 0.000198 Length: 2480
Score: 87.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x US-10-157-031-400 (1-2480)

QY 1 TyrSerGlyProProSerGlyAlaArgArgArgAsnCysTyrGlu 15
|||
DB 850 TACAGCGGCCCCCGAGCGCGCCCGCGCGGAACGTGTACGAA 894

RESULT 2

US-10-723-860-7932
; Sequence 7932, Application US/10723860
; Publication No. US20040233606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; PRIOR FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7932
; LENGTH: 2744
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2534)..(2743)
; OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-7932

Alignment Scores:
Pred. No.: 0.000219 Length: 2744
Score: 87.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x US-10-723-860-7932 (1-2744)

QY 1 TyrSerGlyProProSerGlyAlaArgArgArgAsnCysTyrGlu 15
|||
DB 880 TACAGCGGCCCCCGAGCGCGCCCGCGCGGAACGTGTACGAA 924

RESULT 3

US-10-473-126-45
; Sequence 45, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; FILE REFERENCE:
; TITLE OF INVENTION: proliferative disorders
; CURRENT APPLICATION NUMBER: US/10/473,126
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 45
; LENGTH: 3025
; TYPE: DNA
; ORGANISM: Homo Sapiens

US-10-473-126-45

Alignment Scores:
Pred. No.: 0.000241 Length: 3025
Score: 87.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x US-10-473-126-45 (1-3025)

QY 1 TyrSerGlyProProSerGlyAlaArgArgArgAsnCysTyrGlu 15
|||
DB 1761 TACAGCGGCCCCCGAGCGCGCCCGCGCGGAACGTGTACGAA 1805

RESULT 4

US-10-473-126-183
; Sequence 183, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/473,126
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 183
; LENGTH: 3025
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-183

Alignment Scores:
Pred. No.: 1.55 Length: 3025
Score: 64.00 Matches: 12
Percent Similarity: 80.00% Conservative: 0
Best Local Similarity: 80.00% Mismatches: 3
Query Match: 73.56% Indels: 0
DB: 18 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x US-10-473-126-183 (1-3025)

QY 1 TyrSerGlyProProSerGlyAlaArgArgArgAsnCysTyrGlu 15
|||
DB 1761 TATACGGGTTTTTCGAGCGCGGTTTCGCGCGGAATTGTACGAA 1805

RESULT 5

US-10-473-126-184/c
; Sequence 184, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; FILE REFERENCE:
; TITLE OF INVENTION: proliferative disorders
; CURRENT APPLICATION NUMBER: US/10/473,126
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 184
; LENGTH: 3025
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-184

Alignment Scores:
Pred. No.: 32.7 Length: 3025
Score: 56.00 Matches: 10

```

; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 12548
; LENGTH: 542
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-12548

Alignment Scores:
Pred. No.: 27.2 Length: 542
Score: 52.00 Matches: 8
Percent Similarity: 71.43% Conservative: 2
Best Local Similarity: 57.14% Mismatches: 4
Query Match: 59.77% Indels: 0
DB: 18 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x US-10-357-930-12548 (1-542)
QY 1 TyxSerGlyProSerGlyAlaArgArgAsnCysTyr 14
:::|||||:::|||||
Db 316 CACTCGGGGCCCCCCCCCTCCGAGAGCTCGAAGCTTAT 275

RESULT 8
US-10-282-122A-13812
; Sequence 13812, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Cheryl
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Mi
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13812
; LENGTH: 2358
; TYPE: DNA
; ORGANISM: Burkholderia mallei

```

US-10-282-122A-13812

Alignment Scores:

Pred. No.: 117 Length: 2358
 Score: 52.00 Matches: 9
 Percent Similarity: 91.67% Conservativeness: 2
 Best Local Similarity: 75.00% Mismatches: 1
 Query Match: 59.77% Indels: 0
 DB: 17 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x US-10-282-122A-13812 (1-2358)

QY 2 SerGlyProSerGlyAlaArgArgAsnCys 13

Db 1283 AGCGGGCGCGGAGCGGCTCGCGGAGGAGGACTGC 1318

RESULT 9

US-10-128-714-2252

; Sequence 2252, Application US/10128714

; Publication No. US20030119013A1

; GENERAL INFORMATION:

; APPLICANT: Jiang, Bo

; APPLICANT: Hu, Wengqi

; APPLICANT: Tishkoff, Daniel

; APPLICANT: Zamudio, Carlos

; APPLICANT: Eroszhkin, Alexey M

; APPLICANT: Lemieux, Sebastien M

; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and

; FILE REFERENCE: 10182-018-999

; CURRENT APPLICATION NUMBER: US/10/128,714

; PRIOR FILING DATE: 2002-04-23

; PRIOR APPLICATION NUMBER: US 60/285,697

; PRIOR FILING DATE: 2001-04-23

; PRIOR APPLICATION NUMBER: US 60/287,066

; PRIOR FILING DATE: 2001-04-27

; PRIOR APPLICATION NUMBER: US 60/295,890

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/303,899

; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/316,362

; PRIOR FILING DATE: 2001-08-31

; NUMBER OF SEQ ID NOS: 8603

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2252

; LENGTH: 855

; TYPE: DNA

; ORGANISM: Aspergillus fumigatus

US-10-128-714-2252

Alignment Scores:

Pred. No.: 62.7 Length: 855
 Score: 51.00 Matches: 9
 Percent Similarity: 90.00% Conservativeness: 0
 Best Local Similarity: 90.00% Mismatches: 1
 Query Match: 58.62% Indels: 0
 DB: 15 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x US-10-128-714-2252 (1-855)

QY 4 ProProSerGlyAlaArgArgAsnCys 13

Db 497 CCACCGAGTGGTGGAGAGGAGGAGGACTGC 526

RESULT 10

US-10-128-714-7252

; Sequence 7252, Application US/10128714

; Publication No. US20030119013A1

; GENERAL INFORMATION:

; APPLICANT: Jiang, Bo

; APPLICANT: Hu, Wengqi

; APPLICANT: Tishkoff, Daniel

; APPLICANT: Zamudio, Carlos

; APPLICANT: Eroszhkin, Alexey M
 ; APPLICANT: Lemieux, Sebastien M
 ; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
 ; FILE REFERENCE: 10182-018-999
 ; CURRENT APPLICATION NUMBER: US/10/128,714
 ; PRIOR FILING DATE: 2002-04-23
 ; PRIOR APPLICATION NUMBER: US 60/285,697
 ; PRIOR FILING DATE: 2001-04-23
 ; PRIOR APPLICATION NUMBER: US 60/287,066
 ; PRIOR FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: US 60/295,890
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US 60/303,899
 ; PRIOR FILING DATE: 2001-07-09
 ; PRIOR APPLICATION NUMBER: US 60/316,362
 ; PRIOR FILING DATE: 2001-08-31
 ; NUMBER OF SEQ ID NOS: 8603
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 7252
 ; LENGTH: 855
 ; TYPE: DNA
 ; ORGANISM: Aspergillus fumigatus
 US-10-128-714-7252

Alignment Scores:

Pred. No.: 62.7 Length: 855
 Score: 51.00 Matches: 9
 Percent Similarity: 90.00% Conservativeness: 0
 Best Local Similarity: 90.00% Mismatches: 1
 Query Match: 58.62% Indels: 0
 DB: 15 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x US-10-128-714-7252 (1-855)

QY 4 ProProSerGlyAlaArgArgAsnCys 13

Db 497 CCACCGAGTGGTGGAGAGGAGGAGGACTGC 526

RESULT 11

US-10-128-714-1252

; Sequence 1252, Application US/10128714

; Publication No. US20030119013A1

; GENERAL INFORMATION:

; APPLICANT: Jiang, Bo

; APPLICANT: Hu, Wengqi

; APPLICANT: Tishkoff, Daniel

; APPLICANT: Zamudio, Carlos

; APPLICANT: Eroszhkin, Alexey M

; APPLICANT: Lemieux, Sebastien M

; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and

; FILE REFERENCE: 10182-018-999

; CURRENT APPLICATION NUMBER: US/10/128,714

; PRIOR FILING DATE: 2002-04-23

; PRIOR APPLICATION NUMBER: US 60/285,697

; PRIOR FILING DATE: 2001-04-23

; PRIOR APPLICATION NUMBER: US 60/287,066

; PRIOR FILING DATE: 2001-04-27

; PRIOR APPLICATION NUMBER: US 60/295,890

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/303,899

; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/316,362

; NUMBER OF SEQ ID NOS: 8603

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1252

; LENGTH: 908

; TYPE: DNA

; ORGANISM: Aspergillus fumigatus

US-10-128-714-1252


```
Alignment Scores:
Pred. No.: 66.6 Length: 908
Score: 51.00 Matches: 9
Percent Similarity: 90.00% Conservative: 0
Best Local Similarity: 90.00% Mismatches: 1
Query Match: 58.62% Indels: 0
DB: 15 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x US-10-128-714-1252 (1-908)

Qy 4 ProProSerGlyAlaArgArgAsnCys 13
Db 497 CCACCGAGTGGTGGAGGAAGGAGGAACTGTC 526

RESULT 12
US-10-128-714-6252
; Sequence 6252, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6252
; LENGTH: 908
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-6252

Alignment Scores:
Pred. No.: 66.6 Length: 908
Score: 51.00 Matches: 9
Percent Similarity: 90.00% Conservative: 0
Best Local Similarity: 90.00% Mismatches: 1
Query Match: 58.62% Indels: 0
DB: 15 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x US-10-128-714-6252 (1-908)

Qy 4 ProProSerGlyAlaArgArgAsnCys 13
Db 497 CCACCGAGTGGTGGAGGAAGGAGGAACTGTC 526

RESULT 13
US-10-437-963-48056/c
; Sequence 48056, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
```

```
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 48056
; LENGTH: 2328
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_50768C.1
US-10-437-963-48056

Alignment Scores:
Pred. No.: 170 Length: 2328
Score: 51.00 Matches: 9
Percent Similarity: 75.00% Conservative: 0
Best Local Similarity: 75.00% Mismatches: 3
Query Match: 58.62% Indels: 0
DB: 18 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x US-10-437-963-48056 (1-2328)

Qy 4 ProProSerGlyAlaArgArgAsnCysTyrGlu 15
Db 194 CTTCCATCCGCCGCCGGAAGGAACTGCTAGAA 159

RESULT 14
US-10-128-714-252
; Sequence 252, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 252
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-252

Alignment Scores:
Pred. No.: 212 Length: 2908
Score: 51.00 Matches: 9
Percent Similarity: 90.00% Conservative: 0
Best Local Similarity: 90.00% Mismatches: 1
Query Match: 58.62% Indels: 0
DB: 15 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x US-10-128-714-252 (1-2908)
```

QY 4 ProProSerGlyAlaArgArgAsnCys 13
Db 1497 CCACCGAGTGGTGGAGAGGAGGAACTGC 1526

RESULT 15

US-10-128-714-5252
; Sequence 5252, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5252
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: *Aspergillus fumigatus*
US-10-128-714-5252

Alignment Scores:
Pred. No.: 212 Length: 2908
Score: 51.00 Matches: 9
Percent Similarity: 90.00% Conservative: 0
Best Local Similarity: 90.00% Mismatches: 1
Query Match: 58.62% Indels: 0
DB: 15 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x US-10-128-714-5252 (1-2908)

QY 4 ProProSerGlyAlaArgArgAsnCys 13
Db 1497 CCACCGAGTGGTGGAGAGGAGGAACTGC 1526

Search completed: March 7, 2005, 11:37:51
Job time : 494 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 7, 2005, 06:54:23 ; Search time 3104 Seconds
(without alignments)
183.945 Million cell updates/sec

Title: US-10-018-964-3_COPY_1_15
Perfect score: 87
Sequence: 1 YSGPPSGARRNCYE 15

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10018964/runat_03032005_114707_21373/app_query.fasta_1.199
-DB=EST -QFMT=fasap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -TRANS=human40.cdi -LIST=45
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -MINLEN=15 -MODE=LOCAL
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MAXLEN=2000000000
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10018964 @CGN 1 1 5180 @runat_03032005_114707_21373 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_hcc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gse1.*
9: gb_gse2.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	100.0	580	2	BE296570 601174072
2	87	100.0	621	2	BE295620 601176586
3	87	100.0	678	4	BG829220 602753106
4	87	100.0	958	9	AY399565 Pan trogl
5	87	100.0	963	9	AY399564 Homo sapi
6	87	100.0	1597	3	CB612159 full-len
7	84	96.6	529	8	BZ940212 CH240_107
8	80	92.0	520	2	BZ203994 601869112
9	70	80.5	614	2	BE298779 601119328

10	70	80.5	805	2	BE250705
11	70	80.5	807	2	BF204151
12	70	80.5	833	4	BI115390
13	70	80.5	1055	5	BX359424
14	65	74.7	1000	4	BI114827
15	63	72.4	406	1	AI428972
16	63	72.4	773	9	AY399566
17	63	72.4	964	5	BU511776
18	63	72.4	993	2	BE300478
19	63	72.4	1855	3	AK076157
20	61	70.1	822	4	BG827674
21	60	69.0	702	5	BU443453
22	60	69.0	786	8	CC320860
23	60	69.0	950	5	BU129194
24	60	69.0	1146	8	CC273903
25	58	66.7	830	4	BI455953
26	57	65.5	183	7	CNI68132
27	57	65.5	397	7	CK450832
28	57	65.5	715	7	CK459619
29	57	65.5	721	7	CK459968
30	56	64.4	480	6	CB730730
31	56	64.4	714	6	CD299521
32	55	63.2	500	6	CA604674
33	55	63.2	643	4	BU261450
34	55	63.2	669	4	BU261460
35	55	63.2	687	5	BU646686
36	55	63.2	870	5	BO727321
37	55	63.2	925	4	BI253858
38	54	62.1	510	4	BG905028
39	54	62.1	666	5	BO753201
40	54	62.1	942	9	CL493021
41	53	60.9	406	1	AJ776501
42	53	60.9	553	6	CF064769
43	53	60.9	561	5	BK765940
44	53	60.9	578	5	BX616453
45	53	60.9	621	2	BE260242

ALIGNMENTS

RESULT 1
BE296570
LOCUS 601174072P1 NIH_MGC_17 Homo sapiens CDNA clone IMAGE:3529601 5',
DEFINITION mRNA sequence.
ACCESSION BE296570
VERSION BE296570.1 GI:9180134
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Straubeberg, Ph.D.
Email: csapsb@remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LHC200 row: P column: 18
High quality sequence stop: 579.
Location/Qualifiers
1. 580
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3529601"
/tissue_type="rhabdomyosarcoma"

/lab.host="DH10B (phage-resistant)"

/clone.lib="NIH_MGC_17"

/note="Organ: muscle; Vector: pOTB7; Site:1: EcoRI;

Site 2: XhoI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Size-selected >500bp

for average insert size 1.8kb. Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:

Pred. No.: 0.00434 Length: 580
 Score: 87.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x BE295620 (1-580)

QY 1 TyrSerGlyProProSerGlyAlaArgArgAsnCysTyrGlu 15

Db 326 TACAGCGGCCCCCGAGCGCGCGCGCGGAACTGCTACGAA 370

RESULT 2

BE295620 601176586F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:3531812 5',
 LOCUS mRNA sequence. EST 20-JUL-2000
 DEFINITION 621 bp mRNA linear

ACCESSION BE295620.1 GI:9179172

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCW206 row: 1 column: 21

High quality sequence stop: 607.

FEATURES

Location/Qualifiers

1..621

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3531812"

/tissue_type="rhabdomyosarcoma"

/lab_host="DH10B (phage-resistant)"

/clone.lib="NIH_MGC_17"

/note="Organ: muscle; Vector: pOTB7; Site 1: EcoRI;

Site 2: XhoI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Size-selected >500bp

for average insert size 1.8kb. Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:

Pred. No.: 0.00465 Length: 621

Score: 87.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x BE295620 (1-621)

QY 1 TyrSerGlyProProSerGlyAlaArgArgAsnCysTyrGlu 15

Db 325 TACAGCGGCCCCCGAGCGCGCGCGGAACTGCTACGAA 369

RESULT 3

BE292220 602753106F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4905744 5',
 LOCUS mRNA sequence. EST 23-MAY-2001
 DEFINITION 678 bp mRNA linear

ACCESSION BE292220

VERSION BE292220.1 GI:14176807

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCW1807 row: 1 column: 01

High quality sequence stop: 677.

FEATURES

Location/Qualifiers

1..678

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4905744"

/tissue_type="rhabdomyosarcoma"

/lab_host="DH10B (phage-resistant)"

/clone.lib="NIH_MGC_17"

/note="Organ: muscle; Vector: pOTB7; Site 1: EcoRI;

Site 2: XhoI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Size-selected >500bp

for average insert size 1.8kb. Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:

Pred. No.: 0.00508 Length: 678
 Score: 87.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x BE292220 (1-678)

QY 1 TyrSerGlyProProSerGlyAlaArgArgAsnCysTyrGlu 15

Db 578 TACAGCGGCCCCCGAGCGCGCGCGGAACTGCTACGAA 622

RESULT 4

AY399565
 LOCUS 958 bp DNA linear GSS 15-DEC-2003
 DEFINITION Pan troglodytes MYOD1 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY399565
 VERSION AY399565.1 GI:39755554
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 1 (bases 1 to 958)
 Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 958)
 AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source Location/Qualifiers
 1..958
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..>958
 /gene="MYOD1"
 /locus_tag="HCM0266"

gene

ORIGIN
 Alignment Scores:
 Pred. No.: 0.0072 Length: 958
 Score: 87.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x AY399565 (1-958)

Qy 1 Ty-SerGlyProProSerGlyValaArgArgaenCystvTgLu 15
 |||||
 Db 637 TACAGCGCCCCCGAGCGGCCCGCGCGGAACTGCTACGAA 681

RESULT 5
 AY399564
 LOCUS 963 bp DNA linear GSS 15-DEC-2003
 DEFINITION Homo sapiens MYOD1 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY399564
 VERSION AY399564.1 GI:39755553
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 963)
 Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 963)
 AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source Location/Qualifiers
 1..963
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>963
 /gene="MYOD1"
 /locus_tag="HCM0266"

gene

ORIGIN
 Alignment Scores:
 Pred. No.: 0.00723 Length: 963
 Score: 87.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x AY399564 (1-963)

Qy 1 Ty-SerGlyProProSerGlyValaArgArgaenCystvTgLu 15
 |||||
 Db 637 TACAGCGCCCCCGAGCGGCCCGCGCGGAACTGCTACGAA 681

RESULT 6
 CR612159
 LOCUS 1597 bp mRNA linear HTC 21-JUL-2004
 DEFINITION full-length cDNA clone CS0DI057YK10 of Placenta Cot 25-normalized
 of Homo sapiens (human).
 ACCESSION CR612159
 VERSION CR612159.1 GI:50492966
 KEYWORDS HTC; CNSLT_cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1597)
 Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished
 Contact: Feng Liang Email: fliang@lifetech.com URL: <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
 Faraday Avenue
 2 (bases 1 to 1597)
 Genoscope.
 Direct Submission
 JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 FEATURES
 source Location/Qualifiers
 1..1597
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DI057YK10"
 /tissue_type="Placenta Cot 25-normalized"

```
/plasmid="pCMVSPORT 6"
```

```

/psblastn=pcmvsp01_g"

ORIGIN
Alignment Scores:
Pred. No.: 0.012 Length: 1597
Score: 87.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x CR612159 (1-1597)
QY 1 TyrSerGlyProProSerGlyAlaArgArgAsnCysTyrGlu 15
Db 700 TACAGCGGCCCCCGAGCGCGCGCGCGGCGGAACTGCTACGAA 744

RESULT 7
BZ940212 529 bp DNA linear GSS 13-JUN-2003
LOCUS CH240_107G14.TJ CHORI-240 Bos taurus genomic clone CH240_107G14,
DEFINITION genomic survey sequence.
ACCESSION BZ940212
VERSION BZ940212.1 GI:31725595
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactylia; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
REFERENCE 1 (bases 1 to 529)
AUTHORS Larkin,D.M., Everts-van der Wind,A., Rebeiz,M., Schweitzer,P.,
Bachman,S., Green,S., Campos,E.J., Benson,L.D., Edwards,J., Liu,L.,
Womack,J.E., de Jong,P.J. and Lewin,H.A.
TITLE A Cattle-Human Comparative Map Built with Cattle BAC-ends and Human
Genome Sequence
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: CH240_107G14.TV
Contact: Harris Lewin
Department of Animal Sciences
University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 333 5998
Fax: 217 244 5617
Email: h-lewin@uiuc.edu
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pjejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering/information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by the University of Illinois at Urbana
Champaign, USA with funds provided by grant No. AG202-34480-11828
from USDA-CSREES and AG99-35205-8534 from USDA/NRI (Livestock
Genome Sequencing Initiative)
Plate: 107 row: G column: 14
Seq primer: SP6
Class: BAC ends.
FEATURES
Location/Qualifiers
source
1..529
/organism="Bos taurus"
/mol_type="genomic DNA"
/strains="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_107G14"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/notes="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

ORIGIN
Alignment Scores:
Pred. No.: 0.012 Length: 1597
Score: 87.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x CR612159 (1-1597)
QY 1 TyrSerGlyProProSerGlyAlaArgArgAsnCysTyrGlu 15
Db 700 TACAGCGGCCCCCGAGCGCGCGCGCGGCGGAACTGCTACGAA 744

RESULT 7
BZ940212 529 bp DNA linear GSS 13-JUN-2003
LOCUS CH240_107G14.TJ CHORI-240 Bos taurus genomic clone CH240_107G14,
DEFINITION genomic survey sequence.
ACCESSION BZ940212
VERSION BZ940212.1 GI:31725595
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactylia; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
REFERENCE 1 (bases 1 to 529)
AUTHORS Larkin,D.M., Everts-van der Wind,A., Rebeiz,M., Schweitzer,P.,
Bachman,S., Green,S., Campos,E.J., Benson,L.D., Edwards,J., Liu,L.,
Womack,J.E., de Jong,P.J. and Lewin,H.A.
TITLE A Cattle-Human Comparative Map Built with Cattle BAC-ends and Human
Genome Sequence
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: CH240_107G14.TV
Contact: Harris Lewin
Department of Animal Sciences
University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 333 5998
Fax: 217 244 5617
Email: h-lewin@uiuc.edu
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pjejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering/information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by the University of Illinois at Urbana
Champaign, USA with funds provided by grant No. AG202-34480-11828
from USDA-CSREES and AG99-35205-8534 from USDA/NRI (Livestock
Genome Sequencing Initiative)
Plate: 107 row: G column: 14
Seq primer: SP6
Class: BAC ends.
FEATURES
Location/Qualifiers
source
1..529
/organism="Bos taurus"
/mol_type="genomic DNA"
/strains="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_107G14"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/notes="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

ORIGIN
Alignment Scores:
Pred. No.: 0.012 Length: 1597
Score: 87.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x CR612159 (1-1597)
QY 1 TyrSerGlyProProSerGlyAlaArgArgAsnCysTyrGlu 15
Db 700 TACAGCGGCCCCCGAGCGCGCGCGCGGCGGAACTGCTACGAA 744

RESULT 7
BZ940212 529 bp DNA linear GSS 13-JUN-2003
LOCUS CH240_107G14.TJ CHORI-240 Bos taurus genomic clone CH240_107G14,
DEFINITION genomic survey sequence.
ACCESSION BZ940212
VERSION BZ940212.1 GI:31725595
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactylia; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
REFERENCE 1 (bases 1 to 529)
AUTHORS Larkin,D.M., Everts-van der Wind,A., Rebeiz,M., Schweitzer,P.,
Bachman,S., Green,S., Campos,E.J., Benson,L.D., Edwards,J., Liu,L.,
Womack,J.E., de Jong,P.J. and Lewin,H.A.
TITLE A Cattle-Human Comparative Map Built with Cattle BAC-ends and Human
Genome Sequence
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: CH240_107G14.TV
Contact: Harris Lewin
Department of Animal Sciences
University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 333 5998
Fax: 217 244 5617
Email: h-lewin@uiuc.edu
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pjejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering/information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by the University of Illinois at Urbana
Champaign, USA with funds provided by grant No. AG202-34480-11828
from USDA-CSREES and AG99-35205-8534 from USDA/NRI (Livestock
Genome Sequencing Initiative)
Plate: 107 row: G column: 14
Seq primer: SP6
Class: BAC ends.
FEATURES
Location/Qualifiers
source
1..529
/organism="Bos taurus"
/mol_type="genomic DNA"
/strains="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_107G14"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/notes="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

ORIGIN
Alignment Scores:
Pred. No.: 0.012 Length: 1597
Score: 87.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x CR612159 (1-1597)
QY 1 TyrSerGlyProProSerGlyAlaArgArgAsnCysTyrGlu 15
Db 700 TACAGCGGCCCCCGAGCGCGCGCGCGGCGGAACTGCTACGAA 744

RESULT 7
BZ940212 529 bp DNA linear GSS 13-JUN-2003
LOCUS CH240_107G14.TJ CHORI-240 Bos taurus genomic clone CH240_107G14,
DEFINITION genomic survey sequence.
ACCESSION BZ940212
VERSION BZ940212.1 GI:31725595
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactylia; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
REFERENCE 1 (bases 1 to 529)
AUTHORS Larkin,D.M., Everts-van der Wind,A., Rebeiz,M., Schweitzer,P.,
Bachman,S., Green,S., Campos,E.J., Benson,L.D., Edwards,J., Liu,L.,
Womack,J.E., de Jong,P.J. and Lewin,H.A.
TITLE A Cattle-Human Comparative Map Built with Cattle BAC-ends and Human
Genome Sequence
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: CH240_107G14.TV
Contact: Harris Lewin
Department of Animal Sciences
University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 333 5998
Fax: 217 244 5617
Email: h-lewin@uiuc.edu
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pjejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering/information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by the University of Illinois at Urbana
Champaign, USA with funds provided by grant No. AG202-34480-11828
from USDA-CSREES and AG99-35205-8534 from USDA/NRI (Livestock
Genome Sequencing Initiative)
Plate: 107 row: G column: 14
Seq primer: SP6
Class: BAC ends.
FEATURES
Location/Qualifiers
source
1..529
/organism="Bos taurus"
/mol_type="genomic DNA"
/strains="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_107G14"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/notes="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

ORIGIN
Alignment Scores:
Pred. No.: 0.012 Length: 1597
Score: 87.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query
```


/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:411751"
 /tissue_type="rhabdomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_17"
 /note="Organ: muscle; Vector: pOTB7; Site: 1: EcoRI;
 Site 2: XhoI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:
 Pred. No.: 2.35 Length: 807
 Score: 70.00 Matches: 12
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 80.46% Indels: 0
 DB: 2 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x BF204151 (1-807)

QY 4 ProProSerGlyAlaArgArgAgnCysTyGlu 15
 |||||
 Db 508 CCCCAGCGCGCCCGCGGCGGAACTGCTACGAA 543

RESULT 12

Bi115390

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Bi115390 833 bp mRNA linear EST 26-JUN-2001
 602863245F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:5022419 5',
 mRNA sequence.

Bi115390.1 GI:14566291

EST.

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCMI842 row: 1 column: 12

High quality sequence stop: 769.

Location/Qualifiers

1..833

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5022419"

/tissue_type="rhabdomyosarcoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_17"

/note="Organ: muscle; Vector: pOTB7; Site 1: EcoRI;
 Site 2: XhoI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University

ORIGIN

Alignment Scores:
 Pred. No.: 2.43 Length: 833
 Score: 70.00 Matches: 12
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 80.46% Indels: 0
 DB: 4 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x Bi115390 (1-833)

QY 4 ProProSerGlyAlaArgArgAgnCysTyGlu 15
 |||||
 Db 708 CCCCAGCGCGCCCGCGGCGGAACTGCTACGAA 743

RESULT 13

Bi115390

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Bi115390 1055 bp mRNA linear EST 09-APR-2004
 BX359424 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CSODI057YK10 5-PRIME, mRNA sequence.

Bi115390.2 GI:46291352

EST.

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Li W.B., Gruber C., Jessee J., and Polayes D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On May 5, 2003 this sequence version replaced gi:30372341.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

6204.f

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?c=CSODI057BF05QPI&c=6204.f.

Location/Qualifiers

1..1055

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODI057YK10"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:
 Pred. No.: 3.08 Length: 1055
 Score: 70.00 Matches: 12
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 80.46% Indels: 0
 DB: 5 Gaps: 0

US-10-018-964-3_COPY_1_15 (1-15) x BX359424 (1-1055)

QY 4 ProProSerGlyAlaArgArgAgnCysTyGlu 15
 |||||

THIS PAGE BLANK (USPTO)